

Gore's humanitarianism loses out to strong-arm tactics

Now that the giant drug manufacturers are stamping their feet over proposed compulsory licensing as a means of easing South Africa's AIDS crisis, Al Gore's "values of conscience" appear to have withered away.

Aspiring Democratic presidential candidate Al Gore, who as Vice-President Gore has commonly embraced AIDS-related causes, has been forced to defend himself recently against charges that, on one particularly desperate AIDS issue, he is a cold-hearted lackey of the pharmaceutical industry (see *Nature* 399, 717; 1999).

In South Africa, where antenatal clinic surveys conducted this spring showed 22 per cent of sexually active adults to be infected with HIV, AIDS is expected to slash life expectancy to below 40 years by 2010. The drug cocktails that have curtailed sickness and death in the developed world are unaffordable by all but the most privileged South Africans. Confronting this, the government passed a law in 1997 under which the health minister may authorize local manufacturers to circumvent patents and make far cheaper versions of these drugs, paying only fixed royalties to the rights-holder — so-called compulsory licensing. The law also allows South Africa to import the drugs from countries where drug manufacturers make them available more cheaply.

The international pharmaceutical industry has cried foul, and sued in South Africa, delaying implementation of the law. Meanwhile, Gore and the US Trade Representative have mounted what the US government itself has described as an "assiduous, concerted campaign" to convince South African officials to rewrite or overturn the law.

To be fair, the aggressive language used to depict the US effort was cooked up by the State Department last year only under pressure from congressional Republicans. Led by Rodney Frelinghuysen, a congressman from New Jersey, the home of many giant drug-manufacturers, they had written a foreign-aid bill to withhold aid to South Africa until the Clinton administration convinced them that it was acting on behalf of the drug industry.

Nevertheless, whatever its motivation, the facts of Gore's recent

record remain. Using his role as co-chair, with South African President Thabo Mbeki, of the US–South Africa Binational Commission, the vice-president has repeatedly sought to force a sovereign nation confronting a disastrous epidemic to forswear one of the few means of getting medicine to its desperate people. Coming from a presidential candidate who has been preaching about the importance of "values of conscience" to US political life, this sticks in the gullet.

There are those, including some in the pharmaceutical industry, who argue that South Africa's health infrastructure is insufficiently prepared for these sophisticated drugs. Utilitarian critics also say that the country's finite resources should be used for prevention, not for treating those unfortunate enough to be infected already. Both arguments miss the point. It should not be a case of "either/or", but "both/and". And, indeed, South Africa's move to access state-of-the-art medicines is part of an effort to address the epidemic on all these levels.

There is one final point. Gore and the industry complain that the South African law is in blatant violation of the World Trade Organization (WTO) intellectual-property agreement known as "TRIPS". The agreement allows for compulsory licensing in cases of "national emergency or ... extreme urgency". But there is enough fine print in the agreement, and enough ambiguity in South African law as it is written, to allow honest disagreement on whether South Africa may be in breach of TRIPS.

Here is where the hypocrisy comes in. Instead of using the WTO's well-defined dispute-settlement mechanisms (as it has done recently in disputes with Europe over, for instance, bananas), the United States, led by Gore, has found it convenient to use back-room tactics to try to strong-arm South Africa into changing its ways. This leaves one suspecting that the vice-president's critics are right when they claim that he has been more than a little influenced by the pharmaceutical kings. □

Gems in a millennium's history

Introducing a series of essays celebrating quirks and little-known milestones of the past.

Governed as it is by the most pressing interests and needs of its readers, this publication tends to focus on matters of the present. Happily, our Book Reviews pages often supply a welcome relief from such temporal myopia. In particular, the "In retrospect" feature in that section throws an occasional spotlight on some notable book plucked from the past. The impact of this exercise can be surprising — an essay on the possible influence of a meteorologist on Jane Austen's *Emma* stimulated articles on the front pages of newspapers on more than one continent.

Now a new millennium approaches. The turn of the year 2000 is fundamentally spurious as an anniversary, and special measures can ultimately be justified only by a peculiar human partiality for round numbers — whatever their origin. But the event has turned into something of a stimulus — not only to check computer

systems (which we expect to work just fine), but also to take stock and look back on science's past. Later this year, we shall also look forward.

So it is that, from this week, we apply the principle of "In retrospect" to the wider sweep of science's history. We have invited a number of people — both in science and outside it — to write one-page essays that reflect not necessarily on the great discoveries but on any person or event, drawn from any time in the past 1,000 years, that they feel is worth highlighting for any reason at all — so long as it is in some sense relevant to a scientific readership.

We start with the reflections of Freeman Dyson (see page 27). These weekly contributions will testify not only to the breadth of interest of their authors, but also to the fact that the nooks and crannies of history can be as compelling as its landmarks. □



Sweden sets ethical standards for use of genetic 'biobanks'

[MUNICH] The Swedish Medical Research Council (MRC) last month published Europe's fullest set of ethical guidelines on the use and control of 'biobanks' — tissue and health-science information banks that can be used for genetic research.

The guidelines were developed by the MRC's ethics committee, and their publication follows the founding at the beginning of the year of a private gene-broking company called UmanGenomics, whose *modus operandi* was developed in parallel with the ethics committee's considerations.

UmanGenomics has exclusive rights to generate and sell genetic information from blood samples stored in a 15-year-old medical bank based on samples from 60 per cent of the geographically isolated population of Västerbotten, a county in northern Sweden.

The guidelines, which define biobanks as any banks storing biological samples or information that may be linked to an individual, stress the importance of protecting individual data. Thus, they advise, codes linking data held in a biobank to an individual should be kept within a public institution such as a university or medical authority. These codes should be broken only in exceptional circumstances with the approval of an independent ethics committee.

The ethics committee should also control all access to stored material, and no monopoly on access should be granted. Research projects that might exhaust any part of the stores should not normally be approved.

Informed consent should be sought from donors for every new use of their biological samples. The ethics committee should decide whether or not a research proposal involves a new use, or a use that does not diverge significantly from that for which informed consent has already been received. The quality of the material stored in biobanks, and the rules for selecting material to be stored, should be strictly controlled.

UmanGenomics, based in Västerbotten's county town Umeå (which provided the pun in the company's name), has been accepted by the community because it conforms with these guidelines, says Gisela Dahlquist, professor of paediatrics at the University of Umeå, who chairs the MRC's ethics committee.

The situation is different from that in Iceland, where a population-based biobank is to be generated and placed under the exclusive control of the US-based gene-broking company deCODE (see *Nature* 396, 395; 1998). The initiative has created great concern. According to a recent opinion poll, 25 per cent of Iceland's 250,000 population fear that



Advance guard: pioneering gene-broking company UmanGenomics — 51 per cent owned by local public bodies — will move next year into headquarters in Umeå recently vacated by the army.

deCODE could 'misuse' their genetic information. Moreover, basic researchers complain that they will have to buy information from deCODE at commercial rates.

It was the University of Umeå's new technology transfer unit, called Technology Bridge Foundation, which recognized the unique potential of Umeå's blood bank, now renamed the Medical Bank.

The Medical Bank was set up in 1985 as part of a general health screen after epidemiologists found that the local population had an unusually high mortality from cardiovascular disease. Västerbotten's 260,000 inhabitants were asked by their family doctors to give blood samples every ten years for research into the causes and treatment of the disease. A study in the early 1990s confirmed that there was no bias towards any socioeconomic group in those participating. Such an established and population-based bank, which companies like deCODE are trying to build from scratch, is believed to be unique.

Sune Rosell, former research director at Astra Pharmaceuticals and now temporary chairman of UmanGenomics, says the company worked with the Medical Bank, university researchers and local politicians to create a 'unique formula' for handling ethical issues at different levels.

"There is control at the individual level through informed consent, at the social level through the regional ethics committee which screens all research proposals, and at the population level, since local politicians sit as non-voting members on the boards of both the company and the Medical Bank," he says.

Most important, says Rosell, is the fact that the local university and health authority own a 51 per cent stake in the company, and so have ultimate control over the genetic information that UmanGenomics will make available to pharmaceutical companies.

It is not just sensitivity to ethical issues that underlies the positive local response to the company, he adds. UmanGenomics has promised to bring high-tech jobs to the area by ensuring that research contracts are carried out locally. In addition, university researchers have free access to the Medical Bank, independent of UmanGenomics, which has exclusive commercial rights.

Many pharmaceutical companies have already shown great interest in collaborating with UmanGenomics, says Rosell. Contracts should be signed in the coming months.

In the meantime, four research projects funded by the Technology Bridge Foundation have begun at the university on diseases that have become concentrated in Västerbotten's isolated population. One involves a family that has a high incidence of cardiac infarcts at around 30 years of age, despite normal blood lipid levels. Another involves a family with decreased sensitivity to pain.

Västerbotten's regional ethics committee is responsible for deciding how to handle informed consent for the use of long-archived blood samples for genetic research. This will be arranged on a project-by-project basis, says Dahlquist.

"Informed consent was originally given by donors for cardiovascular research," she says, "so an ethics committee might, for example, require new individual informed consent for a study on schizophrenia. On the other hand — assuming an important, high-quality project addressing a prevalent disease with presumed minimum harm to participants — it might decide a series of newspaper advertisements offering the opportunity to opt out might be appropriate."

The committee will also decide whether informed consent should be secured from relatives whose lives might be affected by the results of a genetic study.

Alison Abbott

Expert group to look at UK cloning law ...

[LONDON] The British government has said that it needs more time before deciding whether to allow cloning techniques — such as the technology that produced Dolly the sheep — to be used on human embryos.

An expert advisory group, chaired by the government's chief medical officer Liam Donaldson, is to be set up to consider the risks and benefits of such research. The group will scour the literature and seek the views of research institutions to establish evidence for potential health benefits of the technology. Its findings will be presented early next year to the government and to the newly formed Human Genetics Commission.

The decision to set up the expert group is a response to calls for the Human Fertilisation and Embryology Act of 1990, which bans the use of cloning techniques on human embryos, to allow what is being described as "therapeutic cloning".

The proposed changes to the law were suggested at the end of last year by the government's Human Genetics Advisory Com-



Donaldson: "we need to proceed carefully."

mission (HGAC) and the Human Fertilisation and Embryology Authority (HFEA), the statutory body responsible for overseeing certain fertility treatments and human embryo research.

"When the 1990 act was passed, the

beneficial therapeutic consequences that could potentially result from human embryo research were not envisaged," the two organizations stated in a recent jointly issued report, *Cloning Issues in Reproduction, Science and Medicine*.

The HFEA/HGAC report calls for cloning techniques to be permitted on human embryos less than 14 days old for the development of treatments for mitochondrial disease, and for treatments for diseased or damaged tissues or organs.

But in its response to parliament last week, the government reiterated its opposition to cloning for reproductive purposes. It said it still needed to be convinced of the need for a change to the law to allow therapeutic cloning.

"It has been suggested that therapeutic cloning techniques might be able to provide immunologically compatible tissue for the treatment of degenerative diseases of the heart, kidneys and cerebral tissue, or repair damage to skin and bone," said Tessa Jowell, the health minister. "We believe that more evidence is required for the need for such research, its potential benefits and risks, and that account should be taken of alternative approaches that might achieve the same ends."

Donaldson says that the government is in no rush to proceed, since it recognizes that there is considerable public concern over issues such as cloning and research on human embryos. "We need to proceed carefully," he says.

He dismisses claims that delays in licensing research on human embryos would drive British scientists to the United States. "I hardly think there'll be a brain drain in the next six months." The government has rejected a suggestion in the HFEA/HGAC report to introduce legislation that specifically outlaws all forms of human reproductive cloning, choosing instead to review the matter in five years' time.

An official from the Department of Health says that the government sees no reason to change the law at present. This is partly because the prospect of successful reproductive human cloning remains uncertain, but also because it believes the existing law is strong enough.

The HFEA/HGAC suggestion was prompted because the law currently prohibits just one type of cloning — the nuclear substitution of any cell while it forms part of an embryo. The technique that was used to create Dolly involves nuclear transfer into an egg, not into an embryo. But both the HFEA and the government believe that the existing law is broad enough to prevent the use of nuclear transfer in an unfertilized egg.

Ehsan Masood

... as US DNA advisory body redefines itself

[WASHINGTON] The US federal committee that debates new gene-therapy protocols is proposing to expand its remit beyond just experiments involving recombinant DNA.

The Recombinant DNA Advisory Committee (RAC) of the National Institutes of Health (NIH) plans to publish a new definition of recombinant DNA this summer in the *Federal Register*, for discussion and possible adoption at a meeting in September.

The change could broaden the committee's charge to reflect a changing and widening field in which an increasing variety of gene-manipulation technologies are challenging traditional notions of genetic engineering.

For instance, technologies are emerging for exploiting RNA — molecules involved in passing on the messages encoded in genes — for purposes such as blocking viral gene transcription. And it is hoped that oligonucleotides, which are tiny stretches of machine-made DNA, may soon be used to correct inherited diseases by substituting for a deficient gene.

A 1976 definition still used by the RAC defines recombinant DNA as "molecules that are constructed outside living cells by joining natural or synthetic DNA segments to DNA molecules that can replicate in a living cell". Experiments that the RAC should monitor, its guidelines say, are those "involving the deliberate transfer of recombinant DNA into human subjects".

"We have not changed the NIH language to keep up to date with new technologies," says Claudia Mickelson, the biosafety officer at the Massachusetts Institute of Technology, who chairs the RAC. The new definition, she says, is part of an effort to make the RAC's guidelines "more accurately reflect the state of molecular biology and genetic engineering in the labs".

A RAC working group that is developing the new definition is working from a far broader 1993 Food and Drug Administration (FDA) policy, which defines gene therapy as "the administration of genetic material in order to modify or manipulate the expression of a gene product, or to alter the biological properties of living cells". It says that gene-transfer experiments include a variety of technologies, ranging from giving patients stem cells modified with a viral vector to using oligonucleotides to correct genetic mutations.

The FDA definition includes cloning, theoretically at least, and is part of the grounds on which the agency asserted its jurisdiction over that area last year (see *Nature* 391, 318; 1998). But Mickelson says that the RAC has no intention of getting involved with cloning issues.

Phil Noguchi, director of the Division of Cellular and Gene Therapies at the FDA, calls the RAC's move to amend the definition "very healthy", adding that "public discussion of these things needs to always be at the cutting edge".

Meredith Wadman

Vast database offers vision of biodiversity

[PARIS] The Organization for Economic Cooperation and Development (OECD) last week approved plans aimed at creating the world's largest biodiversity databank. The Global Biodiversity Information Facility (GBIF) will be launched later this year.

Science ministers from the OECD's 29 member states also decided to redirect its Megascience Forum, a group set up seven years ago to coordinate large scientific projects among member countries. Now called the Global Science Forum, it will emphasize international cooperation to develop global science infrastructures of any dimension.

GBIF is to be launched by an interim steering committee with members from ten countries, with a permanent secretariat to be formed by mid-2000. It will knit together existing databases on biodiversity, to serve as a one-stop information resource. The multi-million dollar project will largely be funded by existing national programmes, such as the effort by the US National Science Foundation (NSF) to digitize natural history data.

Some databases that will contribute to GBIF are already running on websites, including Species 2000, a UK project aimed at indexing all species names. One of GBIF's most important tasks will be to create a comprehensive list of species names, solving problems caused by double references to organisms or misnamed species.

"Sometimes an inappropriate name is used and we find out we've been protecting the wrong organism," says James Edwards, deputy assistant director of the NSF's directorate for biological sciences and chairman



Giving science a hand, clockwise from top left: ministers from Japan, Norway, Russia, South Africa, Israel, Germany, Ireland and Iceland.

of the Megascience Forum bioinformation working group. "Or two countries use different names, so they can't communicate about an endangered species."

GBIF, which is intended to go online in three to four years, will contain scores of databases — including geospatial, chemical, molecular and genetic collections — plus a catalogue of names of known organisms, digitization of natural history data, literature resources, and a bank dedicated to the discovery of new species. It will also include training and outreach programmes to help scientists to use it. Specialists say that GBIF

will change the study of biodiversity.

"Because of distance, it is humanly impossible to get a real vision of biodiversity. A few years down the road we will think 'How could we talk about biodiversity when we couldn't even see it?'," says Frank Bisby, professor of botany at the University of Reading in Britain and chairman of Species 2000.

The idea for GBIF originated three years ago in the Megascience Forum. The new Global Science Forum will tackle such issues as the rivalry between the US Superconducting Collider and Europe's proposed Large Hadron Collider. But it will also consider smaller projects, in the hope of bringing policy-makers and scientists together to resolve international issues.

"We thought, 'Why not extend the mandate so that virtually every project in science can be considered?'," says Michael Osborne, deputy director of the OECD's science, technology and industry directorate. "This will be a significant tool for the scientific community worldwide."

Particle physics and next-generation accelerators could emerge as a hot topic, says Osborne, as could nuclear waste disposal, which the Megascience Forum looked at without drawing up any solutions.

Last week's meeting also decided to form a year-long task force on radioastronomy to sort out conflicts between telecommunications companies and radioastronomers over sharing wavebands (see *Nature* 399, 513; 1999). Radioastronomers have been fighting to prevent mobile phones polluting their designated wavebands. **Heather McCabe**

US 'wastes vital time' as climate-change minority sows confusion

[WASHINGTON] Scientists have accused the US Congress of ignoring solid data on climate change, wasting precious time that should be used to counter the effects of human-generated greenhouse gases.

The group of 57 scientists from 24 states, organized by the Union of Concerned Scientists (UCS), told a press conference this week that the effects of global climate change could be severe for the United States.

They urged Congress to adopt policies to improve energy efficiency and encourage the use of renewable energy. They said Congress should not be distracted by the "contrarian" views of a handful of scientists who dispute what they described as a mainstream scientific consensus on climate change.

"The mainstream scientific community's voices aren't being heard in Congress," says congressman Rush Holt (Democrat, New Jersey), the former director of the Princeton Plasma Physics Laboratory, who took part in the press conference. "It doesn't make

sense to say we need to give equal time to the one or two per cent of scientists who are saying something totally different from the scientific consensus."

The production and consumption of energy is "the number one insult to our globe," says Holt. "[But] I don't think one member out of 50 in Congress would think of it that way." People find it hard to grasp the urgency of an issue mainly reported in complex statistical formats, he says.

Patrick Michaels, of the University of Virginia in Charlottesville, challenges the UCS view. "The bottom line is that the climate models that served as the basis for most of this concern were wrong," says Michaels. "The planet will warm, but not as much as it was feared." He says that government efforts to promote renewable energy and energy efficiency are not needed in a market system in which they will evolve anyway. "Doing very little about this is the best way to solve the problem."

Michaels' views are typical of "minority viewpoints that haven't been adequately peer-reviewed" and that confuse Congress, says Walter Oechel, a professor of biology and director of the Global Change Research Group at San Diego State University. Oechel contributed to the 1995 report of the Intergovernmental Panel on Climate Change, a United Nations-sponsored group of 2,500 scientists who found that human activity was raising atmospheric temperatures.

Last November, the United States signed the Kyoto Protocol, stating a commitment to reduce carbon emissions to seven per cent below 1990 levels between 2008 and 2012. But President Bill Clinton has not submitted the treaty to the Senate for ratification. He says he first wants assurances of meaningful participation from key developing countries such as China and India, and negotiations to be concluded on provisions for emissions trading. These are not expected to be complete until 2001. **Meredith Wadman**

Congress threat to technology transfer

[SAN DIEGO] The transfer of scientific technology from the three national laboratories of the US Department of Energy (DoE) to private companies is being threatened by a congressional proposal for resolving disputes over patent rights.

The proposal would require the three laboratories to offer an arbitration process that could result in substantial damage awards in patent disputes. At the moment, patent disputes are addressed in federal court or by the US Patent and Trademark Office.

Officials of the University of California — which manages the Lawrence Livermore, Los Alamos and Sandia national laboratories for the DoE — say that the proposal could halt technology transfers because the financial risks to a public research laboratory would become too high.

University officials see no reason why the dispute procedures should change. They say that the proposal is a politically driven attempt by a small Alabama company to try to win the rights to a potentially valuable radar technology patented by a former Livermore scientist. The new technology, called micro-impulse radar (MIR), has wide application from sensor devices to wireless communication. The Alabama company, Time Domain Corporation of Huntsville, claims it has a right to MIR because it owns the patent from which this technology derives.

MIR technology is one of several discoveries licensed by Livermore to private firms as part of a nascent technology transfer programme worth more than \$5 million in annual licence fees.

But as the programme has grown, it has come under attack from at least two private companies which are developing products with technology similar to that licensed by Livermore. The companies complain that the government-funded laboratories are competing against them.

In one recent instance, a Californian company sought congressional assistance to win an agreement that would have blocked Livermore from developing a new type of battery with a zinc-air fuel cell. "It is inappropriate to do that," says Martin Simpson, a university attorney. "The new technology is a public asset that should be developed for the public benefit."

Ronald Cochran, Livermore's executive director, says that the energy department labs have never before experienced a dispute like the one currently involving Time

Domain, which has tried to use political pressure to muscle MIR technology from the federal laboratory.

Records and interviews show that Time Domain employs a former top DoE administrator to lobby the government. The company has funnelled more than \$7,600 into the coffers of Congressman Robert E. 'Bud' Cramer (Democrat, Alabama) and caused the US Patent and Trademark Office to reel under political pressure.

Ralph Petroff, Time Domain's president, says that his company, which has no products at the moment, is fighting for the MIR technology because it is derived from a US patent granted to the firm's scientific director, Larry Fullerton.

But electronics engineer Thomas E. McEwan, who holds more than 20 US patents associated with the MIR technology that he developed at Livermore before leaving the laboratory in 1996, says that Time Domain's contention is "absurd".

McEwan, who was named Distinguished Inventor of the Year in 1995 by the Intellectual Property Owners Association for his MIR work, said that his work differs fundamentally from that of Fullerton.

In an effort to resolve the dispute, Livermore asked the patent office to re-examine McEwan's MIR patent. But McEwan and university officials say the patent office has since come under political pressure from Time Domain advocates, something that the patent office denies.

At times, the battle has turned nasty. Time Domain and its advocates have hinted at a plot to deprive the company of its technology. McEwan says that Time Domain has repeatedly disseminated misrepresentations to smear him and win influence.

Time Domain could have sued in federal court to assert patent infringement, but chose not to because it would cost too much, says Petroff. But McEwan says the company is trying to use politics to win a battle it could not hope to win in court.

The proposal in Congress says the three laboratories must permit an aggrieved private company to seek arbitration, either binding or non-binding, in a patent or licensing dispute. If the company wins, it could, according to the proposal, demand damages and costs, which could fall on the DoE or the University of California.

Late last month, the House of Representatives passed a massive defence-spending bill that included an amendment for the national laboratory arbitration proposal. That measure will now be considered by a joint conference with the Senate in September. Another amendment with similar provisions for dispute resolution is to be voted on by the House early this month. **Rex Dalton**



Japan to launch ambitious genome project

[TOKYO] Japan is to launch a national programme to identify and map single nucleotide polymorphisms (SNPs) in the Japanese population, according to a report released last week by the government's *ad hoc* working group on SNPs research.

The project will aim to map between 100,000 and 150,000 SNPs in two years, using samples obtained from at least 50 Japanese people. Data from other Asian populations will be added to the study at a later date.

SNPs represent the most abundant type of variation in the human genome, and can be used to map genes involved in common diseases such as diabetes and coronary heart disease.

Maps of this kind can also be used to identify individuals who are likely to be unresponsive to, or suffer side effects from, various drugs — allowing their doctors to choose the safest and most effective form of treatment for them.

The planned project will be carried out jointly by government ministries, universities and industry. It will create a public SNP database, which scientists hope will contribute to international efforts such as the SNP Consortium. This consists of 10 major pharmaceutical companies and five leading publicly funded genomics institutes, none Japanese (see *Nature* 398, 545; 1999).

Unlike programmes led by the SNP Consortium and companies such as Genset, which are based on genomic SNPs, the Japanese project will use SNPs in coding sequences by analysing full-length complementary DNA and expressed sequence tags.

But critics question Japan's ability to meet its projected target, given its relative weakness in information processing technology and sequencing capacity, and given the lack of Japanese firms that would be able to compete with Western biotechnology companies. **Asako Saegusa**

Reactor may switch to low-grade uranium

[MUNICH] Germany's controversial new FRMII research reactor, being built to run on weapons-grade highly enriched uranium (HEU), could be converted to use low enriched uranium (LEU) by 2008, says a new report.

The DM810 million (US\$430 million) 20 megawatt neutron source would not need to be completely rebuilt, and the conversion would not damage relevant scientific research, according to the expert committee's report, published last week.

The Technical University of Munich (TUM) has continued building the reactor, currently scheduled for operation by 2001, despite protests that the use of HEU in research reactors breaks international proliferation agreements. The new German Social Democrat and Green coalition government committed itself to investigate the possibility of converting the reactor in its coalition agreement.

The report outlines three possibilities for the FRMII, but makes no recommendation. The first — immediate conversion to the use of existing LEU fuel, along with a power upgrade to 32MW to maintain neutron flow density — is certain to be ruled out on cost grounds.

The second option — immediate conversion to the use of LEU without increasing the reactor's power — would delay its operation by at least three years. The expert committee considers this a "reasonable" period, but it would also result in a 20–30 per cent reduction of the neutron flow density. The university says the costs of this option would still be very high, at about DM300 million.

The third option — building and operating the reactor as planned and converting it to use a specially developed low or medium enriched fuel element, geared to FRMII's design, when it becomes available — would cost only about DM12 million, says the report. Moreover, the reduction of the neutron flow density would be only seven per cent, which "would cause no considerable limitation of the scientific use of the FRMII".

The United States and France have industrial research programmes to develop suitable LEU fuel for research reactors. The report concludes that if Germany and the university are serious about converting FRMII, they should take part in this research.

But the expert committee estimates that such fuel would take around seven years to develop. During this time, proliferation problems related to the interim use of HEU must be taken seriously. It warns that continued use of HEU could weaken international efforts to prevent the fuel being used for weapons, and says that "less trustworthy" countries could cite FRMII to justify their own HEU reactor projects.

The university has welcomed the "precise



Building work: the FRMII reactor is currently being built to use weapons-grade uranium.

technical survey" which the report provides, but rejects its political statements. "Addressing the issue of proliferation was not among the commission's tasks", says Gert von Hasel, spokesman for the FRMII project.

But Hans-Josef Fell, the Green parliamentary group's spokesman on research, says the university's view of FRMII as a purely scientific matter is "politically unacceptable".

The Greens are aware of the "central importance" of neutron research, says Fell, but the immediate conversion of the reactor is a high political priority among party leaders. "We must not be ignored once again by

our SPD coalition partner," he says, referring to the Greens' previous concessions to the SPD over nuclear power policies (see *Nature* 397, 189; 1999).

But conversion of the reactor seems likely to be delayed. The university, supported by the Bavarian government, has said it would agree to a future conversion, but for now insists on building and operating the neutron source according to the initial design.

The options will be discussed between federal government and the Bavarian state government. A decision is expected within the next six months.

Quirin Schiermeier

European Union tightens GMO regulations

[PARIS] Regulations on the environmental release of genetically modified organisms (GMOs) were tightened last week by ministers of the 15 member states of the European Union. Meeting in Luxembourg, they voted that GMOs should be given only provisional marketing approvals for ten years and not unlimited authorizations as at present. France, Ireland and Italy abstained.

Under the agreement, initial approvals would be considered only for releases "meeting certain safety criteria where sufficient experience exists". They would also require an environmental risk assessment, be subject to "extensive consultation" with scientists and the public, and meet a series of compulsory monitoring requirements. Products would also need to be labelled clearly as containing GMOs.

A moratorium on new approvals, proposed by Greece and supported by France, was rejected. The new regulations are unlikely to come into force for at least

two years, as they require the approval of the European parliament. This may therefore amount to a *de facto* moratorium.

Paul Muys, a spokesman for EuropaBio, an association that represents the European biotechnology industry, argues that the fact that the European Union has agreed on a common position is progress. At the same time, he deplores what he sees as unnecessary new layers of regulation for a technology that poses no new risks and is already tightly regulated.

A requirement to renew approvals after ten years is unnecessary, argues Muys, as existing regulations allow for products to be withdrawn if evidence of risks emerges. He adds that a non-binding statement issued after the meeting by France, Denmark, Italy, Greece and Luxembourg, calling for a freeze on approvals until the new regulations come into force, confirms industry's worst fear: that a moratorium on GMOs is now in place in Europe.

Heather McCabe & Declan Butler



The full text of items on this page — and further details and background information about the Unesco/ICSU World Conference on Science, which has been taking place in Budapest this week — can be found on *Nature's* website at <http://www.nature.com>

Academies call for ban on patenting agricultural life forms

The Council of the Third World Academy of Sciences (TWAS), which met in Budapest last Thursday (24 June) before the start of the main conference, endorsed a call for a ban on the patenting of all “agricultural life forms”.

“Agriculture in much of the developing world is the result of the collective experience gained from the sweat and toil of poor peasants over thousands of years,” says TWAS vice-president Muhammad Akhtar, emeritus professor of biochemistry at the University of Southampton in England.

“Recombinant DNA technology is unlikely to alter more than a fraction of one per cent of the existing genomic make-up of edible plants,” he says. “It is hence an affront to the sense of fairness that multinational corporations should be able to claim the ownership, through patenting, of living systems, for such a minuscule contribution.”

Earlier in the week, a meeting of the scientific advisory body to the United Nations biodiversity convention in Montreal had given a lukewarm response to a more farmer-friendly alternative to ‘Terminator’ seeds, the controversial technology in which seeds are genetically modified to become sterile after one season’s planting (see *Nature* 399, 721; 1999 & helix.nature.com/wcs).

The alternative is known as T-Gurt — Trait-specific Genetic Use Restriction Technology. T-Gurt seeds are modified to produce specific traits, such as tolerance to drought. The trait is activated after the seed is sprayed with a proprietary chemical. The seed will still germinate without the chemical, but will not have the modified characteristics. **E.M.**

Debt relief cash ‘should be spent on a fund for science’

The Africa group of countries has agreed on what could be the conference’s most innovative proposal so far: creating a science fund for poor countries from the debt relief that was agreed last month by the G8 nations, meeting in Cologne, for the 40 least developed countries.

The proposal was agreed by science and education ministers from Africa at a meeting on Monday (28 June) in Budapest. It will be put to the conference drafting group by Senegal, which represents Africa on the group.

Under the proposal, countries with heavy burdens of debt will be urged to agree to set up a science fund from money earmarked for debt repayments. The repayments will be waived under the plan agreed by the industrialized countries that met in Cologne.

Support for the idea could also come from the 50-member Organization of Islamic Countries (OIC), whose ministers also met on Monday (see below). At this meeting, Atta ur Rahman, coordinator-general of the OIC’s Ministerial Standing Committee on Scientific and Technological Cooperation (Comstech), proposed the setting up of a separate intergovernmental fund in which countries would be repayed — for example through the commercialization of research — according to the size of their contributions.

Ministers from Africa also agreed that the draft conference documents need to put additional emphasis on two other issues. The first is “that science is the common heritage of mankind, whose results and benefits should be shared equitably”. The second is the need for a more clearly defined mechanism to follow up implementation of the



Nlend: ‘science must serve all humanity’

conference resolutions among Unesco’s member countries.

“The idea that some countries are producers of science while others are consumers cannot be allowed to continue,” says Hogbe Nlend, Cameroon’s minister for science.

The fund idea has emerged unexpectedly. Many, if not most, conference delegates had come to Budapest expecting little tangible outcome from the meeting. Indeed, some important developed countries have said that they will resist firmly any calls for funds.

This is partly a result of the experience of the last major United Nations science conference, in Vienna in 1979, where a battle between rich and poor countries over a science-for-development fund ended in stalemate. That experience is also believed to be a reason why there is no mention of funds in any of the draft conference documents.

“The draft agenda only deals with [funding] vaguely,” says Nlend. “But for poor countries, the lack of funds is a critical issue.” This is echoed by Rahman of Comstech who said that “without a fund, nothing will happen”.

Nlend reveals that officials from African and OIC countries studied the lessons from Vienna, one of which was not to push too hard for a fund that relies largely on donations from developed countries, but to explore ways in which developing countries can help themselves. **Ehsan Masood**

Islamic nations set themselves tough target for science budgets

Science and education ministers of the Organization of Islamic Countries (OIC) endorsed a draft strategy for developing science in the Muslim world on Monday (28 June). The main, and most controversial, recommendation is that governments should allocate a minimum of one per cent of their gross domestic product for scientific and technological development.

But many of the ministers admitted that a goal of one per cent was unrealistic. Delegate after delegate said that such a target would be too difficult for most OIC member countries to achieve.

The strategy document is the work of the Islamic Educational, Scientific and Cultural Organization and the OIC’s Ministerial Standing Committee on Scientific and Technological Cooperation (Comstech).

Atta ur Rahman, Comstech coordinator-general, said that all developing countries had no option but to increase their investment in science if they want to lift themselves above the pile of underdeveloped nations. A pressing priority, he said, is human resource development. A second priority is regional cooperation.

“Forgive me for being blunt,” Rahman, a

chemist, told the ministers. “I am a working scientist, not a diplomat, and am not blessed with the appropriate words with which to decorate this message: too often, meetings like this result in very nice-looking resolutions. But enough has been said, it is time to act.” Rahman will receive this year’s Unesco science prize at a ceremony today.

Abd al Rahman al Awadi, of the Islamic Organization for Medical Sciences in Kuwait, called on the ministers to emphasize the “moral and ethical dimension” to science in their responses to the draft documents of the World Conference on Science. **E.M.**

Clinton's adviser foresees a new breed of civic scientist

Researchers must be "global citizen scientists" — aware of both new knowledge and the needs of society — according to Neal Lane, science adviser to US President Bill Clinton.

In a speech delivered on his behalf by Bruce Alberts, president of the National Academy of Sciences, Lane said that "in their new capacity of 'civic scientist', scientists and engineers must step outside of their campuses, laboratories and institutes to engage in an active dialogue with their fellow citizens".

Scientists need to look beyond intriguing research questions and examine the ways in which new scientific knowledge may be best used by society, Lane said.

This, he added, means that scientists need to communicate with — and listen to — the public. "There is a great need for the public to have a better understanding of science. But there is an equally great need for scientists to have a better understanding of the public." □

Young scientists seek help from Unesco

Young scientists have called on Unesco to create a body to represent their views. The International Forum of Young Scientists (IFYS) wants a forum for discussing general issues and challenges to science.

The group feels that a permanent body would keep young scientists in contact with both Unesco and the International Council for Science. Unesco director-general Federico Mayor, who met with an IFYS delegation, reacted positively to the request and said the prospects for such a body were "very good".

"This is one of the things I wish the World Conference on Science to take into account. I think it is a good initiative and I have told young scientists that we must try to have a permanent interaction."

Mayor says he will "consider whether a young people's forum could be a permanent representative in follow-ups to the World Conference.

"The young people should not be silent and should be making some noise." He envisaged interaction at regional and subregional levels via the Internet.

More than 150 young scientists and

students from 54 countries met at the three-day forum organized by Unesco to let the "new generation" of scientists contribute to the conference. Their final recommendations include suggestions that: scientists try harder to inform the public about research and its implications; science education at all levels be strengthened; education presents science in a cross-disciplinary manner; ethical aspects be considered in all scientific undertakings and education programmes; and that scientists take more responsibility to help less developed countries.

Nobel laureate Leon Lederman warned against treating young scientists as "cheap labour", saying "we often talk about the responsibility of scientists towards society ... a big part of our responsibility is to see how young scientists develop". Lederman urged scientists to communicate the importance of science funding as well as the need for science literacy.

Meanwhile, a Unesco survey of young scientists found that their main concern was over the correct use of science to achieve peace and development. **Natasha Loder**

Research 'must aid fight against poverty'

M. S. Swaminathan, one of the architects of the Green Revolution, has called for the eradication of poverty to be given the highest priority by national governments and international organizations.

In a keynote address on the opening day of the World Conference on Science, Swaminathan said that poverty eradication is the key to meeting basic human needs. "Poverty is the root cause of hunger, lack of shelter and access to clean water, illiteracy, ill health and other forms of human deprivation," he said.

He showed how many applications of modern science, its management and its administration, have contributed to poverty. But he said that there is little doubt that science has a crucial role to play in poverty eradication — both in the way it is applied and in the way it is used to empower the poor to help themselves.

An example of the first, he said, is the genetic modification of agriculture. An example of the second is the way science enables farmers to manage land and livestock better by training them to gather and make use of computer-based data on soil, pests and water. This helps them to understand the causes of specific problems as well as to think of ways of resolving them.

In a wide-ranging address, Swaminathan pointed out that no international summit or convention in recent times has addressed the

issue of poverty exclusively. He added that the goals of many such gatherings will not be reached without simultaneously addressing poverty issues.

Three such issues, according to Swaminathan, are: establishing an international centre for cooperation in water management; a global programme to fight maternal and fetal malnutrition to cut the number of children born with low birth weights; and a scheme to eliminate hidden hunger caused by the absence of micronutrients in the diet.

Contaminated drinking water contributes to ten per cent of diseases in developing countries. And in 1996, 1.8 billion low- and middle-income households lacked access to sanitary facilities. But, with global per capita water supplies 30 per cent lower than 25 years ago, and continuing to decline, Swaminathan predicted that conflicts over water would grow unless active measures were taken to resolve potential disputes.

Similar measures are needed to tackle malnutrition in many countries, according to Swaminathan. Up to half of the children in several developing countries have low birth weight because their mothers are undernourished, he said. "Millions of children will continue to be born for mere existence and not for happiness if this area of nutrition continues to receive inadequate attention."

But poverty, said Swaminathan, had

another cause: lack of access to knowledge, or what he described as "orphans remaining orphans". One example, he said, was the "technological apartheid" being created by the increasing amount of science being covered by intellectual property rights.

This is closing off sources of knowledge to the poor, and is inequitable. For example, the lack of recognition by the World Trade Organization (WTO) of the "invaluable contributions of tribal and rural families to genetic resources conservation" remains a source of concern. He called on the WTO's member states to correct this by bringing the Trade Related Intellectual Property Rights agreement in line with the United Nations Convention on Biological Diversity.

The world's science academies, said Swaminathan, have a duty to help the rural poor. "A priority task for the Inter-Academy Center proposed by Bruce Alberts [president of the US National Academy of Sciences] should be the closing of the vast knowledge and skill gap between rich and poor nations and between the rich and poor within nations."

Any anti-poverty policy, said Swaminathan, had to emphasize access to science for the women of the developing world. Some of the blame, he suggested, lay at the door of science education, which tended to be biased towards men. **E.M.**

Senior energy official quits in wake of US secrets scandal

[WASHINGTON] The head of the US Department of Energy's nuclear weapons programme last week became the first official to leave his job in the fall-out from allegations that the Chinese stole nuclear secrets from the Los Alamos National Laboratory in New Mexico.

Victor Reis, Assistant Secretary of Energy for Defense Programs, submitted his resignation to Bill Richardson, the energy secretary. The *Washington Post* reported that Reis disagreed with Richardson about how to improve security at the weapons labs, favouring putting them in the charge of a semi-autonomous nuclear agency. Reis refused to comment on why he was leaving.

The Clinton administration has admitted that senior White House officials were told of China's possible theft of secrets in July 1995, a year earlier than originally contended.

Cloning giant pandas comes a step closer

[TOKYO] Scientists from China's National Academy of Sciences announced last week that they have made a significant step

towards cloning the country's national symbol, the giant panda.

According to a report by China's Xinhua News Agency, a team led by Chen Dayuan, professor of zoology at the academy, succeeded in producing cloned hybrid embryos by transferring the nucleus of an adult giant panda cell into an egg obtained from a rabbit. The scientists are now attempting to implant the embryo into the uterus of a foster mother of a second species, thought to be a black bear.

This is part of a national programme aimed at cloning the giant panda in a bid to save the animal from extinction (see *Nature* 394, 409; 1998). But other scientists have voiced doubts about the experiment's chance of success, saying that cross-species cloning needs to be carried out using genetic material from closely related species.

German grants agency defends funding policy

[MUNICH] Germany's main university funding agency, the Deutsche Forschungsgemeinschaft (DFG), wants to retain its traditional 'bottom up' funding policy, despite an external evaluation committee's call for more technology foresight.

"Any additional control of the direction of

research, even when initiated by scientists, would damage the autonomy of German science," said Ernst-Ludwig Winnacker, DFG president, at the agency's annual assembly last week.

The evaluation committee had recommended last month that the DFG should develop strategic research programmes on its own initiative to help increase the competitiveness of research in Germany (see *Nature* 399, 395; 1999). But Winnacker said that the agency is already supporting particularly promising areas of research, for instance bioinformatics, in specially designed programmes. Additional 'steering elements' would inappropriately favour a top-down approach in the DFG's research funding procedures, he said.

Trademark bid shatters the peace of café society

[PARIS] A storm has broken out among the chattering classes of Parisian scientific society over a decision by the French Physics Society (SFP) to apply for a trademark on the concept of 'Bar des Sciences' — informal meetings between scientists and the public held in cafés. The idea is based on the famous French 'cafés philo'.

The initiative has enjoyed spectacular success since its launch in 1997 and has

spread from Paris to the provinces. The SFP originated the idea, but as many as half the 'Bar des Sciences' are organized by other researchers. Many science journalists say they find the idea of a trademark outrageous. "If we amuse ourselves by patenting or trademarking colloquiums, seminars or other forms of scientific meeting, it is going to be a farce," says Marc Mennessier, editor of *Le Journal du CNRS*.

Gilles Bogart of the SFP says he is surprised by the strength of the reaction. Bogart defends the society's action as an innocent bid to protect the name from being exploited by pseudo-scientific groups, sects and commercial companies.

European patents on up and up as fees fall

[MUNICH] The number of patent applications filed at the European Patent Office (EPO) in Munich increased by 13 per cent in 1998, reaching a record high of 113,000.

In presenting the 1998 EPO annual report last month, EPO president Ingo Kober said that the significant increase is mainly a consequence of the organization's policy of slashing registration fees (see *Nature* 393, 722; 1998). The fee, currently US\$770, is to be reduced further in July by 18 per cent.

Leaders in the number of patent

applications submitted to the EPO are, by a wide margin, the United States (28.6 per cent), Germany (19.6 per cent) and Japan (16.8 per cent) — followed by France (6.9 per cent) and Great Britain (4.8 per cent).

Medical and veterinary technologies stay top of the list of subjects patented (6,800 applications). Electronic communication technologies were number two (6,200 applications) and showed the highest increase (22 per cent). Applications in biochemistry and genetic engineering increased by 16 per cent and rank eighth.

US public 'confused' about environment crises

[WASHINGTON] The US public is confused by what it sees as scientific disagreement on environmental threats such as global warming, according to an analysis of public opinion commissioned by the American Geophysical Union.

"What the public needs to hear from the scientific community is a greater sense of agreement about what is known," wrote John Immerwahr, associate vice-president of Villanova University, who conducted the study for the nonprofit organization Public Agenda.

Based on focus group discussions in five US cities and a review of recent poll results,

the study also found that the public accepts the seriousness of the global warming problem, but doubts society's ability to solve it. The underlying cause is seen as "a pervasive climate of selfishness and greed". Because environmental problems are seen as politically intractable, "communication efforts designed to 'pump up the volume' on the seriousness of geophysical problems may backfire".

Science and education escape German cuts

[MUNICH] Science and education have been exempted from a decision by the German cabinet to crop 7.4 per cent across the board from the general government budget for 2000. The science budget will remain at its 1999 level of DM15 billion (US\$8 billion).

The Max Planck Society and the Deutsche Forschungsgemeinschaft, Germany's universities granting agency, will find their budgets increased by three per cent. This is good news and bad news. The federal government had originally promised a five per cent increase, but a few weeks ago the regional government finance ministers said they would only support a two per cent increase (see *Nature* 399, 508; 1999).

The budget has to be approved by the German parliament in September.

Weathering a storm of global statistics

Sir— Another El Niño 'event of the century' has come and gone, and with it a wide variety of estimates of how much damage the world has incurred, ranging from US\$14 billion to \$69 billion¹⁻³. Damage from all natural disasters in 1998 alone is believed to have cost some \$93 billion⁴. Of greater concern than this lack of accuracy is a general tendency to describe impacts through a few global totals.

By aggregating impacts into a single estimate, much of the knowledge of and insight into the human-climate interface is lost. The environment does not affect us in simple, one-number packages. Only recently have we begun to measure quantitatively the varied losses and benefits (impacts) associated with climate variability, and these estimates have yet to be referenced to a baseline of impacts occurring in 'normal' years⁵.

Regional variations in reported damages, for example, do not inherently represent the reality of losses. Burton *et al.*⁶ note that biases exist "toward

overestimating losses from industrialized countries and underestimating losses in developing countries or in areas remote from centers of government and mass media". It is not easy to attribute and distinguish between losses associated with climate variability and maintenance costs. Although biases and uncertainties do exist, other factors need to be accounted for, for example the use of a prejudiced vocabulary to describe losses. Many estimators are unfortunately tempted to represent a globally perceived value rather than inter-regional functionality, for example. A thatched house may not fetch a global market value, but its loss during a flood is nonetheless an impediment to livelihood maintenance or to development.

Climate impacts are also largely a function of perception and scale. What is devastating to an individual is not likely to register on an international scale, except in extreme calamities. The incentives, disincentives and cultural preferences of an affected area often guide this perception of

impact, which influences the reporting and visibility of climate events. It is likely that incentives such as 'relief aid' have increased the number of reported events. There are many other complex variables that need to be taken into account.

Not all impacts associated with climate variability are disastrous or even negative. To better understand the interaction between humans and climate, the popular vocabulary used to describe and record impacts of climate variability must move beyond deaths and dollar losses.

Kelly Sponberg

Office of Global Programs, National Oceanic and Atmospheric Administration, 1100 Wayne Avenue, Suite 1210, Silver Spring, Maryland 20910, USA

1. Obassi, G. O. P. Secretary-General, WMO, Presentation at International Seminar on the 1997-98 El Niño Event: Evaluation and Projections, Guayaquil, Ecuador (9 November 1998).
2. McPhaden, M. J. *Nature* **398**, 559-562 (1999).
3. NOAA Office of Global Programs *Compendium of Climate Variability* Draft (May 1999).
4. Munich Re Group Press Release (15 March 1999).
5. Glantz, M. *Climate Change* **1**, 305-306 (1978).
6. Burton, I., Kates, R. & White, G. *The Environment as Hazard* 2nd edn (Guilford, New York, 1993).

Colonial adventures

Sir— Walter Gratzer, in his review of Freeman Dyson's book, writes: "Dyson's guess, based on a typical interval between discovery of a new land (America in 1492) and its settlement by outsiders (the arrival of the *Mayflower*), is that colonization of space could begin in about 2085" (*Nature* **398**, 770; 1999). This perpetuates a mistaken view of the early settlement of the Americas. The first lasting European colony in the Americas was established by Nicolás de Ovando in Hispaniola in the West Indies in 1502, only ten years after Christopher Columbus's first voyage and more than a century before the *Mayflower*.

Michael A. G. Michaud

11 Chemin des Tulipiers, 1208 Geneva, Switzerland

Modern museums are far from fossilized

Sir— In his review of Steven Conn's book *Museums and American Intellectual Life, 1876-1926*, Thomas Gieryn asserts that long ago "museums ceased to be centres of intellectual work" (*Nature* **399**, 31-32; 1999). He offers reasons for this decline, which, among other things, "made natural history collections . . . almost irrelevant to the production of new knowledge".

While these and similar comments may accurately portray the changing role of many museums earlier this century, they fail to acknowledge the predominant and vital role that is played by both public and private research museums in contemporary intellectual life. This is especially true in the United States, where several prominent universities (such as Berkeley, Harvard, Kansas and Michigan) have wisely continued to support, and even expand, their research museums.

The result of this investment is that many such museums now constitute the centres of intellectual enquiry in several fields, including evolutionary biology, anthropology and human biology. Knowledge of the past offers many important lessons, but it is not an infallible guide to the present or future.

James Hanken

Museum of Comparative Zoology, Harvard University, Cambridge, Massachusetts 02138, USA

Secondary metabolism and the risks of GMOs

Sir— The potential problems of altering the chemical composition of crops were discussed in your Briefing¹. One aspect of this debate relates to secondary metabolism, which is an attractive area to exploit because of the importance of such

compounds in resistance, defence and product quality. In our view, the rules governing the evolution and role of secondary metabolites need to be discussed and understood in order to understand the risks associated with the genetic modification of crops.

We have previously proposed^{2,3} a model based on the well-known fact that potent, specific biological activity is a rare property for a molecule to possess. In this model, organisms with a rich secondary metabolism (most plants and many microbes) have gained fitness by possessing metabolic traits that enhance the production and retention of chemical diversity. Two such traits could be a broad substrate tolerance of some of the enzymes involved in secondary metabolism, and the utilization of matrix pathways.

Two examples from terpenoid metabolism illustrate the metabolic flexibility proposed. In spearmint (*Mentha gracilis*), a mutation caused an enzyme to produce a new product, but several other new compounds were also made, at least one of which was unpredictable⁴. In the grand fir (*Abies grandis*), two enzymes can make multiple products from a single substrate (one produces 52 and the other 34)⁵. If such metabolic properties exist in all organisms with a rich secondary metabolism, the introduction of a gene could potentially have quite unpredictable outcomes, as in the following examples.

First, the introduction of an enzyme

expected to produce a single new chemical could also produce other new compounds owing to the substrate tolerance of existing enzymes. Second, the introduction into a new organism of a gene encoding an enzyme involved in secondary metabolism could produce more than one product owing to the substrate tolerance of the introduced enzyme.

Third, the introduction of a gene into an organism could disturb secondary metabolite production simply as a consequence of the random gene insertion, with unplanned and unexpected increases in the content of some compounds, owing to changes in the metabolic flux through matrix pathways.

In the Briefing¹ it was suggested that metabolite profiling or clinical trials might help address the issue of the unknown consequences of manipulating food composition. Both these approaches might be very helpful when assessing the consequences of introducing a single major product, but they would be less productive when assessing the consequences of interfering with secondary metabolism.

Of major concern is the fact that the secondary metabolite profiles of plants can vary considerably, so the effect of introducing a gene into a plant might be predictable only under defined conditions that may not be achievable in the field. The secondary metabolite profile is complex, and extremely small amounts of highly potent compounds can have profound biological consequences — how complete would metabolite profiling have to be?

Richard D. Firn*, **Clive G. Jones†**

*Department of Biology, University of York, York YO1 5DD, UK

†Institute of Ecosystem Studies, Millbrook, New York 12545, USA

1. Butler, D. *et al.* *Nature* **398**, 651–656 (1999).
2. Jones, C. G. & Firn, R. D. *Phil. Trans. R. Soc. Lond.* **333**, 273–280 (1991).
3. Firn, R. D. & Jones, C. G. *Rec. Adv. Phytochem.* **30**, 295–312 (1996).
4. Croteau, R. *et al.* *Plant Phys.* **96**, 744–752 (1991).
5. Steele, C. L. *et al.* *J. Biol. Chem.* **273**, 2078–2089 (1998).

Making sense of GM tomatoes

Sir — Your Briefing on genetically modified (GM) crops¹ refers to Zeneca's GM tomato puree. Hans-Jörg Buhk is reported as saying that “the best performing line was caused by an unpredicted ‘sense’ event (gene activation). This was a rare event, either a contamination or a chance turnaround [in the genome]”. This is wrong.

The phenomenon of sense down-regulation by short sense constructs was discovered in earlier research². It has nothing to do with gene activation. Subsequently, for commercial

development, genetic modification of a processing cultivar of tomato was carried out with a short sense construct; 210 individual transgenics were produced and these formed the basis for further selection, breeding and development. All the development work has been rigorously reviewed in the United States and United Kingdom by the regulatory bodies. It is wrong to describe the development of this product as being due to “contamination or a chance turnaround”. It is misleading to claim this as evidence that GM products are somehow unpredictable.

You show a picture of a can of GM tomato paste with a prominent label. The product was labelled voluntarily in line with our policy of openness, as has been widely reported. Labelling was not a requirement for commercialization. The caption “Backlash: consumer concerns forced action on labelling” is clearly misleading.

Our target was not ‘shelf life’ as this is a product sold in a can.

Simon Bright, Wolfgang Schuch

Zeneca Plant Science, Jealott's Hill Research Station, Bracknell RG42 6ET, UK

1. Butler, D. *Nature* **398**, 651–656 (1999).
2. Smith *et al.* *Mol. Gen. Genet.* **224**, 477–481 (1990)

Winners and losers in Framework programme

Sir — We have examined the performance of the countries included in the recently completed fourth Framework programme of research and technology in the European Union (EU). We found that this programme cost about 3% of total research and development (R&D) expenditure within the EU, but exceeded 6% of the national expenditure of countries that devote a small share (less than 1%) of their gross domestic product (GDP) to R&D (Greece, Portugal and Spain).

Austria, Italy, France and Germany were net donors of funds, particularly Germany, whose scientists received only 18% of the funds for their country's 30% contribution to the budget (Fig. 1). Most other countries were net recipients of funds, with excess funds received by the United Kingdom and Greece representing 30% and 140% of their contributions, respectively.

The success rate for grant applications was similar among all countries at 27% of the proposals filed, except for the remarkable 39% success rate of UK scientists. Hence the roles of the member states as net donors or recipients of funds largely depended on the involvement of their scientists in the programme. This varied, in per capita terms, 10-fold between Germany, where scientists filed the fewest proposals, and Greece.

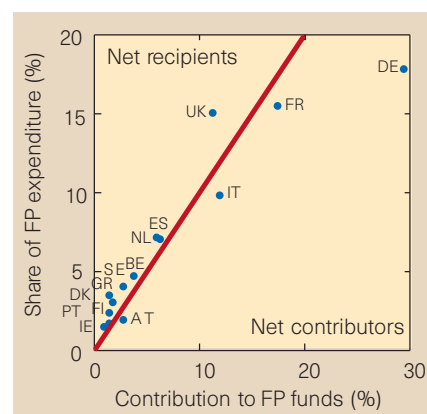


Figure 1 Relationship between contributions and receipts of EU member states participating in the fourth Framework programme (FP). Participants: Austria (AT), Belgium (BE), Germany (DE), Denmark (DK), Spain (ES), Finland (FI), France (FR), Greece (GR), Ireland (IE), Italy (IT), The Netherlands (NL), Portugal (PT), Sweden (SE), and United Kingdom (UK). Data derived from information on funding decisions by the programme's committees. Data are not available for Luxembourg and non-member participants.

Proposals to the Framework programme require the coordination of multinational research groups and the preparation of thorough management plans, which often outweigh the scientific or technical merits of the proposal in the evaluation process. These difficulties deter researchers who can obtain funds from less complex sources. The resources provided by the programme are far more attractive to scientists from countries where R&D resources are scarce (Greece, Portugal and Spain) than to those in countries where R&D investment is comparatively abundant (for example, Germany and France).

The substantial economic benefits the R&D systems of the EU countries with the lowest per capita GDP obtained from their participation in the programme should be of general benefit to the EU. The development of a stronger R&D capacity in countries where this sector has been relatively weak should deliver tangible mid-term benefits to the entire partnership at a time when countries' economies are linked through a common currency.

The fifth Framework programme (1999–2004) will improve the EU's R&D capacity further if the application procedure is simplified and the member states allocate more resources to assist scientists, thereby achieving greater participation and overall quality.

Carlos M. Duarte*, **Maria-José Delgado†**

*Instituto Mediterráneo de Estudios Avanzados, CSIC-Univ. Illes Balears, Carretera de Valldemossa km 7.5, 07071 Palma de Mallorca (Islas Baleares), Spain
†Subdirección General de Proyectos Científico Técnicos, Ministerio de Educación y Cultura, C/ Rosario Pino, 14-16, 28020 Madrid, Spain

Why Africa needs agricultural biotech

There is urgent need for the development and use of agricultural biotechnology in Africa to help to counter famine, environmental degradation and poverty. Africa must enthusiastically join the biotechnology revolution.

Florence Wambugu

The public debate on transgenic crops in Europe is centred on fear and mistrust, quite possibly resulting from the experience over 'mad cow disease'. A recent report¹ from the Food Safety Authority of Ireland to address European Union concerns on genetically modified (GM) crops concluded that there is no evidence that transgenic foods are unsafe. The report, by a group led by Patrick Wall, the authority's chief executive, says that concern in Europe is based on ethical, socio-economic and anti-multinational issues; lack of knowledge or misinformation; environmentalism; food labelling; and consideration of the needs of developing countries.

Many of these concerns have nothing to do with food safety. Transgenic foods are eaten daily in the United States, Australia, Canada, Mexico and elsewhere with no reported undue effects²⁻⁴. Nevertheless, the experts' advice does not seem to influence public opinion in Europe, probably because of a strong anti-biotechnology lobby that actively promotes misinformation and fear, and also because in some cases people have had good reason to distrust 'expert' pronouncements.

One example of Europeans' concern for the Third World is 'terminator technology' — plants engineered to be sterile. But this technology is only a concept and is not being further developed. No products are planned for Africa or elsewhere. Critics of biotechnology have used the fear of this technology to promote serious anti-multinational attitudes — for example, crops in trials have been burnt in some parts of the world.

Another concern promoted by critics of food biotechnology is that of toxins or allergies. An example is the case of the unpublished study by Arpad Pusztai, formerly of the Rowett Research Institute in Scotland, who suggested that rats fed with GM potatoes expressing a snowdrop lectin were slowly being poisoned. After an independent scientific review, these results were found to be misleading and to have been misinterpreted⁵. But the anti-biotechnology lobby is still using them strongly to advance its case in Europe, even though transgenic foods are rigorously tested for possible toxins and allergens before commercialization.

Surely there are parallels to be drawn with an antibiotic such as penicillin, which has continued to be used for many years despite many people being allergic to it because the benefits clearly outweigh the risks. Why is the same reasoning not applied to transgenic foods, where risks at even this low level are not



Biotech could boost productivity: the average maize yield in Africa is less than half the global average.

proven? The anti-biotech lobby also cites as controversial the recombinant DNA processes used to develop transgenic foods. But the same processes are used to develop numerous pharmaceuticals for humans and animals, and many other industrial products. The public seems prepared to accept the application of GM techniques to new pharmaceutical products but not to food production. Why should there be different standards for crops and pharmaceuticals, particularly in Africa where the need for food is crucial for survival?

African perspective

The critics of biotechnology claim that Africa has no chance to benefit from biotechnology, and that Africa will only be a dumping ground or will be exploited by multinationals^{6,7}. On the contrary, small-scale farmers in Africa have benefited by using hybrid seeds from local and multinational companies, and transgenic seeds in effect are simply an added-value improvement to these hybrids. Local farmers are benefiting from tissue-culture technologies for banana, sugar cane, pyrethrum, cassava and other crops. There is every reason to believe they will also benefit from the crop-protection transgenic technologies in the pipeline for banana, such as sigatoka, the disease-resistant transgenic variety now ready for field trials. Virus- and pest-resistant transgenic sugar-cane technologies are being developed in countries such as Mauritius, South Africa and Egypt.

The African continent, more than any other, urgently needs agricultural biotechnology, including transgenic crops, to improve food production. African countries need to think and operate as stakeholders, rather than

accepting the 'victim mentality' created in Europe. Africa has the local germplasm, some of it well-characterized and clean, being held in gene banks in trust by centres run by the Consultative Group of International Agricultural Research. It also has the indigenous knowledge, local field ecosystems for product development, capacities and infrastructure required by foreign multinational companies.

The needs of Africa and Europe are different. Europe has surplus food and has never experienced hunger, mass starvation and death on the regular scale we sadly witness in Africa. The priority of Africa is to feed her people with safe foods and to sustain agricultural production and the environment.

Africa missed the green revolution, which helped Asia and Latin America achieve self-sufficiency in food production. Africa cannot afford to be excluded or to miss another major global 'technological revolution'. It must join the biotechnology endeavour. Transgenic food production increased from 4 million to 70 million acres worldwide from 1996 to 1998 with measurable economic gains and with sustainable agricultural production². It would be a much higher risk for Africa to ignore agricultural biotechnology. Africa's crop production per unit area of land is the lowest in the world. For example the production of sweet potato, a staple crop, is 6 tonnes per hectare compared to the global average of 14 tonnes per hectare. China produces on average 18 tonnes per hectare, three times the African average. There is the potential to double African production if viral diseases are controlled using transgenic technology.

The African continent imports at least 25 per cent of its grain. The use of biotechnology

to increase local grain production is far preferable to this dependence on other countries, particularly as the population growth rate exceeds food production. The inability to produce adequate food forces Africa to rely on food aid from industrialized nations when mass starvation occurs. Although biotechnology is not the only answer to this problem, Africa should certainly benefit in many ways from its use, for example in improved seed quality and resistance to pests and diseases.

The average maize yield in Africa is about 1.7 tonnes per hectare compared to a global average of 4 tonnes per hectare. Some biotechnology applications can be used to reduce this gap, for example in the case of the maize streak virus (MSV), which causes losses of 100 per cent of the crop in many parts of the continent. A biotechnology-transfer project is under way to develop MSV-resistant varieties. The project is brokered by the International Service for the Acquisition of Agri-Biotech Applications (ISAAA), and involves the collaboration of the Kenya Agricultural Research Institute (KARI), the University of Cape Town, the International Centre for Insect Physiology and Ecology in Kenya, and the John Innes Centre in the United Kingdom. Funding is coming from the US Rockefeller Foundation, and Novartis in Europe has donated some technology to KARI.

Researchers at KARI are studying the mechanism of MSV resistance and trying to map the genes responsible. Advanced biotechnology skills, including the use of advanced agroinoculation techniques and molecular markers, is at the core of this effort. A priority in Kenya is also to produce high-yielding, drought-tolerant crop vari-

eties to boost food production in the 71 per cent of the country that is arid or semi-arid.

Africa needs biotechnology to solve its environmental problems, and there is unlimited public demand for agricultural biotechnology products and services. In Kenya, the demand for tree seedlings reaches 14 million per year, whereas the country can only supply 3 million, a clear indication of the need for tissue-culture and cloning techniques to curb deforestation and boost reforestation using indigenous species threatened with extinction. These technologies are being successfully used in South Africa, and ISAAA has facilitated a project for application in Kenya. There are issues of intellectual property rights and patents that require hard work to develop or acquire, and advanced agricultural biotechnology skills will be needed. There may also be a need to work out collaboration agreements with the private sector or with companies that already have patents.

Biotechnology in Africa is needs-based. After working at KARI for nearly a decade to help improve sweet-potato production using traditional breeding and agronomy methods, I made no progress. An opportunity to work in the private biotechnology sector abroad resulted in the development of a transgenic variety that is resistant to sweet-potato feathery mottle virus, which can reduce yields by 20–80 per cent. Control of this disease will improve household food security for millions. This project involved collaboration between KARI, a project called Agricultural Biotechnology for Sustainable Productivity, funded by the US Agency for International Development, and Monsanto. The work by Kenyan scientists focuses on local varieties, and there will be a smooth and sustainable transfer of the technology, which will be shared with neighbouring countries. Kenyan scientists have been trained in gene technology techniques. ISAAA has been asked to help with the transfer and licensing agreement. Similar projects are under way for bananas, sugar cane and tropical fruits.

Remaining problems

Needless to say, Africa has many problems — a shortage of skilled people (especially in biotechnology), poor funding of research, lack of appropriate policies and civil strife. Nevertheless, countries such as South Africa, Egypt, Zimbabwe and Kenya are taking practical steps to ensure that they can use biotechnology for sustainable development.

African countries need to avoid exploitation and to participate as stakeholders in the transgenic biotechnology business. They need the right policies and agencies, such as operational biosafety regulatory agencies, breeders' rights and an effective local public and private sector, to interface with multinational companies that already have the technologies. Consumers need to be informed of the pros and cons of various agricultural

biotechnology packages, the dangers of using unsuitable foreign germplasm, and how to avoid the loss of local germplasm and to maintain local diversity. Other checks and balances are required to avoid patenting local germplasm and innovations by multinationals; to ensure policies on intellectual property rights and to avoid unfair competition; to prevent the monopoly buying of local seed companies; and to prevent the exploitation of local consumers and companies by foreign multinationals. Field trials need to be done locally, in Africa, to establish environmental safety under tropical conditions.

The main goal is to find a balanced formula for how local institutions can participate in transgenic product development and share the benefits, risks and profits of the technology, as they own the local germplasm needed by the multinationals for sustainable commercialization. New varieties must not simply replace local ones. The removal of genes that were in the public domain into the private sector raises concern in Africa.

All these issues mean that Africa must strengthen its capacity to deal with various aspects of biotechnology, including issues of biosafety, creating and sustaining gene banks, and encouraging the emergence of a local biotechnology private sector. The great potential of biotechnology to increase agriculture in Africa lies in its 'packaged technology in the seed', which ensures technology benefits without changing local cultural practices. In the past, many foreign donors funded high-input projects, which have failed to be sustainable because they have failed to address social and economic issues such as changes in cultural practice. The criticism of agribiotech products in Europe is based on socioeconomic issues and not food safety issues, and no evidence so far justifies the opinion of some in Europe that Africa should be excluded from transgenic crops. Africans can speak for themselves. □

*Florence Wambugu is director of the International Service for the Acquisition of Agri-Biotech Applications (ISAAA AfriCenter), Regional African Office, PO Box 25171, Nairobi, Kenya.
e-mail: f.wambugu@cgiar.org*

1. www.fsai.ie
2. James, C. *Global Review of Commercialized Transgenic Crops* ISAAA Brief No. 8 (ISAAA, 1998).
3. Food and Agriculture Organization *Biotechnology and Food Safety* Food and Nutrition Paper No. 61 (FAO, Rome, 1996).
4. *Genetically Engineered Food Production Gathers Pace* CSIRO Media Release 99/117 (CSIRO, Guelph, Canada, 1999).
5. Loder, N. *Nature* **399**, 188 (1999).
6. Christian Aid and other groups in *GM-FREE: Keeping your Life and Environment Free of Genetically Modified Food* Vol. 1, No. 2 (KHI, Skelmersdale, June/July 1999).
7. Nuffield Council on Bioethics *Genetically Modified Crops*, www.nuffield.org/bioethics/publication/modifiedcrops/index.html

The ISAAA receives about 60% of its funding from philanthropic foundations such as Rockefeller, McKnight and Hitachi; 30% from bilateral agencies such as Danida (Denmark), BMZ (Germany) and USAID; and 10% from biotech companies such as Novartis, Agrevo, Pioneer and Monsanto. The views expressed here are the author's and do not necessarily represent those of ISAAA.

GUY MANSFIELD / PANOS PICTURES



Break the culture of dependency: Africa is forced to import at least 25 per cent of its grain.

Taking stock of new flavours

Harold McGee

What's in a taste? Chemists can isolate the components of flavours, and biologists have begun to explain how our bodies enjoy them, but cookery will remain more art than science for a while yet.

Thirty years ago, in a celebrated Friday Evening Discourse at the Royal Institution entitled "The Physicist in the Kitchen", Nicholas Kurti deplored the scientific neglect of cooking, an "insufficiently dignified" activity that nevertheless nourishes and gives daily pleasure to much of humankind. "I think it is a sad reflection on our civilization that while we can and do measure the temperature in the atmosphere of Venus, we do not know what goes on inside our soufflés."¹ In May, at the fourth meeting* of a gastrophysical workshop founded by Kurti and named in his memory, scientists and chefs gathered to explore what goes on inside food and its consumers to generate the all-important sensation of flavour.

The workshop (directed by H. This, *Pour la Science*, Paris) follows an unusual format of Kurti's devising: open discussion among basic scientists, food scientists from industry and universities, and professional cooks, seasoned with brief lectures and informal experiments. The agenda is to analyse and help advance the fine art of cooking as it is practised in domestic and restaurant kitchens.

Flavour is the quality that most often distinguishes excellent from ordinary foods. The summit of the chef's art is to conceive and realize a multicourse meal that progresses through a series of flavours without repetition (S. Hill, *The Merchant House*, Ludlow, Shropshire). However, the flavour of even the simplest dish presents a tremendous challenge to scientific analysis. A given food contains hundreds or thousands of chemicals that stimulate either the tongue's taste receptors or the nose's olfactory receptors, some in parts per billion. Diners have different receptor ensembles and different neural circuitries for reporting and integrating their outputs. Much of the workshop was devoted to sampling these chemical and biological complexities.

An essential culinary preparation is the meat stock, a water extract of meat, bones, vegetables and herbs that is often concentrated by boiling off much of its water — and presumably many of the volatile compounds that contribute to aroma. Why is the reduction often more flavourful, not less? Boiling



Figure 1 *The Cook*, an engraving by Hubert François Gravelot. The verse below reads "New cooking every year, because every year tastes change; and every day there are new ragouts; so be a chemist, Justine". Courtesy of the Museum of Fine Arts, Boston, Sargent Collection.

drives off some flavours but generates others. A portion of veal stock was boiled down to one-third of its original volume, both the stock and its rediluted reduction were analysed by gas chromatography, and then a similar stock and rediluted reduction were served to participants (A. Blake, Firmenich, Geneva). Both stock and reduction were full-flavoured, but the prominent vegetal aromas in the stock were replaced in the reduction by a less easily dissected intensity. The complex gas-chromatograph trace of the reduction showed a loss of several plant-derived volatiles (cinnamaldehyde, eugenol, γ -terpinene, humulene) but augmented di- and trisulphides, probably from the shallots, which reinforce the important sulphurous components of meat flavour. Different flavour balances result when stock is concen-

trated by freezing or by successive additions of fresh extractables, as in the *marmite perpétuelle*, or immortal stew. Volatiles lost during the boiling of stocks and jams could be retrieved with a simple (but probably illegal) distillation apparatus.

Pure samples of individual volatiles were presented for sniffing, many of them suggestive of whole foods. Benzaldehyde conjured cherry; eugenol, clove; γ -terpinene, carrot; *z*-nonenal, cucumber; and 3-*N*-butylidene phthalide, both celery and walnut (F. Benzi, Firmenich, Geneva). Chemists suggested that chefs could use such concentrates and extracts the way a painter mixes colours on a palette, quickly fine-tuning a dish with no need for the original flavouring materials or their lengthy preparation. For example, a drop of hexanal, the fugitive 'green' note in many vegetables and fruits, restores the impression of freshness in a cooked dish. Chefs replied that such a prospect is intriguing, but extracts are not yet a match for good fresh ingredients, and it is easier to modulate flavour with handfuls than serial dilutions. Freshness can be restored to asparagus soup by adding a purée of the raw tips just before serving (F. Blank, Deux Cheminées, Philadelphia). Participants readily distinguished two experimental ratatouilles pitting fresh thyme and bay laurel against their extracts, probably by differences in both flavour qualities and concentrations (A. M. De Gennes, Le Boudin Sauvage, Orsay).

Flavour perception is a dynamic process, especially in the eating of raw foods whose composition undergoes rapid change. Mass spectrometry capable of time-resolved analysis of nostril air flow (50 samples per breath) reveals that the flavour of raw tomato evolves during the course of chewing as various enzyme systems are activated by tissue disruption (R. Linforth, Univ. Nottingham). Endogenous aromatics (isobutylthiazole) reach the nose first, followed after about 30 seconds by hexanal and other oxidation products of unsaturated fatty acids, and then by the corresponding alcohols. Savouring food thus both prolongs the sensation of flavour and enriches it.

Interactive effects complicate the sensation of taste and odour mixtures. The bitterness of quinine in bittersweet tonic water was abolished by a small dose of sodium; in an aroma mixture dominated by cinnamaldehyde (cinnamon), the initially imperceptible vanillin (vanilla) emerged once the cinnamaldehyde receptors had adapted to their high dose (G. Beauchamp, Monell Center, Philadelphia). Taste sensations can also influence aroma perception. In a model system of chewing gum, sugar and mint volatiles, mass spectrometry shows that the actual nostril concentration of volatiles remains high during chewing, whereas the impression of mintiness declines along with the washing-out of sugar (either slowly or

*International Workshop on Molecular Gastronomy N. Kurti, Erice, Sicily, 6–10 May 1999.

rapidly depending on the gum design). This is why children learn to revive spent gum by rolling it in the sugar bowl (UK participants).

Thanks to genetic and developmental differences in receptors and associated neural circuitry, individual humans inhabit distinct flavour worlds. These remain largely unexplored, and their mapping will be furthered by the new Centre Européen des Sciences du Goût at Dijon (C. Masson, CNRS). The inability to detect an aroma can sometimes be overcome through repeated exposure. This has been shown for androsthenone, a steroidal volatile present in truffles, pork and celery, which proved to be aroma-less for half the participants (including a Michelin three-star chef), and either pleasantly earthy or disgustingly rank to the rest. Although general consensus recognizes the involvement of several hundred distinct receptors in the olfactory system and just five in taste, some controversial discrimination studies suggest a taste space of ten or more dimensions (A. Faurion, Ecole Pratique des Hautes Etudes, Massy). Following the recent cloning of putative mammalian taste receptors², chefs should soon find out whether they will

have undreamt-of tastes to work with.

In a French engraving of 1759, a young woman at the stove is told that times and tastes are changing; every day there are new stews: "... so be a chemist, Justine" (Fig. 1; ref. 3). Professional cooks did become inventively systematic, and the result was what we now call classic French cuisine. Today, chemistry is changing even faster than tastes and stews, and seems likely to accelerate their evolution. Will this give twenty-first-century gastronomers another great cuisine, of unprecedented subtlety and diversity? Or will it mean the triumph of receptor-tickling virtual food? Or both at once?

One whiff: because they both contain traces of barnyard indole, a flavourist urged chefs to try combining jasmine flowers and pork liver. Certainly there will be interesting puddings to prove! □

Harold McGee is at 838 La Jennifer Way, Palo Alto, California 94306, USA. He is the author of *On Food and Cooking: The Science and Lore of the Kitchen*, and *The Curious Cook* (Scribner, Collier).

e-mail: hjmcgee@alummi.caltech.edu

1. Kurti, N. *Proc. R. Inst. G. Br.* **42**, 451–467 (1969).

2. Hoon, M. A. et al. *Cell* **96**, 541–551 (1999).

3. Wheaton, B. *Savoring the Past* 198 (Univ. Pennsylvania Press, Philadelphia, 1983).

Materials science

A century of zero expansion

Peter Mohn

On page 46 of this issue¹ Mark van Schilfgaarde, Igor Abrikosov and Börje Johansson present a quantum-mechanical explanation for the strange behaviour of a special class of 'Invar' alloys, so called because of their invariable volume when heated. The experimental finding dates back to 1897 when the Swiss physicist Charles Édouard Guillaume² found that face-centred cubic (f.c.c.) alloys of iron (Fe) and nickel (Ni) with a composition of roughly 35% Ni and 65% Fe exhibit almost zero thermal expansion over a broad temperature range. His discovery immediately found widespread application in the construction of calibrated, high-precision mechanical instruments, such as seismographs. He also developed 'Elinvar', a material that has a negligible change in elasticity when heated, and which for decades was used to make springs for mechanical watches. Guillaume won a Nobel Prize in Physics in 1920 for the discovery of these ferronickel alloys.

Today, Invar alloys are used in many temperature-sensitive devices, such as surveying tapes, and perhaps most notably in shadow masks for television and computer screens. The shadow mask prevents the outer edges of the electron beams from hitting the wrong phosphor dot on the screen. Any distortion caused by heat from the beam would disturb the positioning, producing loss of colour

purity and vertical definition. This can most easily be avoided by manufacturing the mask from Invar (Fig. 1). To celebrate a century of Invar research, a symposium was held in 1997 to review the experimental and theoretical results³.

The thermal expansion of matter comes

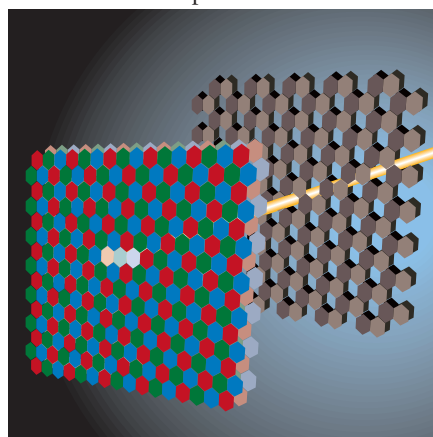


Figure 1 Invar shadow mask. Shadow masks behind computer screens are essential for directing the electron beam to the correct phosphor dot. An Invar alloy, such as $\text{Fe}_{65}\text{Ni}_{35}$, is the material of choice for mask manufacturers, because its dimensions change very little with temperature. This 100-year-old effect is now explained in microscopic detail by van Schilfgaarde et al.¹.

about because of the increased vibration amplitudes of atoms around their positions in the crystal lattice, as the temperature rises. At room temperature the thermal expansion coefficient, α , is roughly constant and usually ranges from 10×10^{-6} to $20 \times 10^{-6} \text{ K}^{-1}$. All Invar alloys are magnetic, so that α is only almost zero below the magnetic-ordering temperature T_C (the Curie temperature) and rises to a higher value for temperatures above T_C . This behaviour made it clear that magnetism must be entangled in the explanation of Invar behaviour.

The first theory pointing in this direction came from Richard Weiss, who proposed the so-called 2γ -state model⁴. He assumed that there were two different magnetically ordered states, a ferromagnetic ground state γ_1 (with magnetic spins aligned parallel) having a larger volume, and an antiferromagnetic state γ_2 (with magnetic spins aligned antiparallel) with a smaller volume, but at a slightly higher energy. As the temperature rises, the γ_2 -state becomes thermally excited, so that the smaller volume of this state compensates for the ever-present vibrational thermal expansion. This leads to the low thermal-expansion coefficient. Above the magnetic-ordering temperature T_C , where the system becomes paramagnetic (that is, with magnetic spins randomly oriented), no further magnetic compensation for the thermal expansion is possible, so that α catches up with the usual high value described above.

Although the 2γ -state model gives a satisfactory explanation for the Invar phenomenon, the microscopic nature of these two states, if they actually exist in the proposed way, remains unclear. It was not until the advent of quantum-mechanical calculations of the electronic and magnetic structure of metallic solids that further insight became possible. During the past two decades there have been a number of promising attempts to illuminate these magnetic states. What makes this new study so exceptionally different is the fact that the authors have included non-collinear spin ordering in their simulations (that is, spins may be tilted at arbitrary angles to each other if such a configuration lowers the energy of the overall spin system). Their calculation shows (see their Fig. 1 on page 47) that at large volumes (corresponding to low temperature) the ground state is indeed given by parallel (ferromagnetic) spin alignment. When they reduce the volume (thus simulating increasing temperature) the spins gradually depart from parallel alignment so that the spin directions become increasingly disordered. To fully understand the enormity of this result one must recall that, in all magnetic materials, increasing the temperature leads to greater disorder of the spin alignment (in this sense, the magnetic-ordering temperature T_C is the temperature where the disorder is complete, so that no net magnetic moment of

the sample remains). For the Invar alloys this increasing disorder is volume dependent and so leads to a decreasing volume, thereby compensating for the vibrational thermal expansion.

This study provides a microscopic explanation for an effect that has puzzled scientists for more than a century. It also paves the way for future work on this class of materials. As well as the Invar effect there is the anti-Invar effect, in which an exceptionally large thermal expansion is found. This effect could also be used profitably, for example to make actuators for micromachines. It is also recognized that all Invar alloys have compositions that are surprisingly close to having

structural instabilities. For example, in Fe–Ni alloys with a nickel concentration below 32%, the crystal structure is no longer f.c.c. but body-centred cubic. It is likely that these issues will keep solid-state physicists busy for years to come. □

Peter Mohn is at the Center for Computational Materials Science, Vienna University of Technology, Getreidemarkt 9/158, A-1060 Vienna, Austria.
e-mail: phm@cms.tuwien.ac.at

1. van Schilfgaarde, M., Abrikosov, I. A. & Johansson, B. *Nature* **400**, 46–49 (1999).
2. Guillaume, C. E. C. *R. Acad. Sci.* **125**, 235–238 (1897).
3. Wassermann, E. F. in *The Invar Effect: Centennial Symposium* (ed. Wittenauer, J.) 51 (Minerals, Metals & Materials Society, 1997).
4. Weiss, R. *Proc. R. Soc. Lond.* **82**, 281–288 (1963).

Cell cycle

A snip separates sisters

Andrew Murray

A paper by Uhlmann *et al.*¹ on page 37 of this issue sheds light on one of the mysteries of cell division — how sister chromosomes separate from one another as cells divide. To transform itself into two genetically identical daughters, a cell must perform three reactions on its chromosomes (Fig. 1a). The sequence begins at a stage known as interphase, where each chromosome replicates to generate a pair of sister chromosomes. Next, cells enter mitosis (nuclear division), and the two sisters align to face in opposite directions on the segregation machinery, or mitotic spindle. Finally, at the end of mitosis, the sister chromosomes separate and move to opposite poles of the spindle.

This chromosomal dance depends on the physical linkage between the two sisters, which allows them to align correctly, but must then be dissolved so they can separate at the end of mitosis. Logic suggests that the linkage must be established during DNA replication and then persist until mitosis; the sisters are uniquely close to each other during the act of replication, making it easy to link one sister to another rather than to a non-sister. If the sisters were not linked during replication, this spatial distinction would rapidly disappear as they diffused apart from each other.

Looking at extreme cases reveals that the stability of sister linkage can vary over seven orders of magnitude. In a human oocyte, for example, the sisters must remain linked to each other from the end of DNA replication until the oocyte enters the meiotic divisions that will transform it into an unfertilized egg — an interval that can exceed 50 years. Failure of this linkage is an important cause of Down's syndrome. By contrast, during early divisions of the fruitfly embryo, the chromosomes must receive the signal to separate and dissolve their linkage in less than two minutes.

Until recently, two obstacles limited our understanding of sister-chromosome linkage: ignorance about how sister separation is triggered at the right stage of the cell cycle, and our inability to see individual chromosomes in yeasts (which have dominated genetic analysis of the cell cycle). The first obstacle was removed with the discovery that cleavage of mitotic cyclins is required for exit from mitosis. The enzyme that targets cyclins for cleavage, the anaphase-promoting complex, also triggers the destruction of unknown proteins, which hold sister chromosomes (also known as sister chromatids) together, possibly by acting as a physical glue

between them². The second obstacle disappeared with the development of methods to locate specific sequences in yeast nuclei, first by *in situ* hybridization on fixed cells^{3,4} and, later, by localization of fluorescent DNA-binding proteins to segments of DNA in living cells⁵.

With these tools in hand, geneticists looked for mutants that perturb sister-chromosome linkage and separation. They identified a protein called Pds1 in budding yeast, or Cut2 in fission yeast, that had to be destroyed by the anaphase-promoting complex for the sisters to separate^{6,7}. But Cut2/Pds1 does not bind to chromosomes, so it cannot be the glue that holds the sisters together. Instead, it controls the activity of another enzyme (Cut1 in fission yeast or Esp1 in budding yeast), which seems to cause separation of sister chromosomes. Initially Cut2/Pds1 helps to get Cut1/Esp1 into a form that will ultimately allow it to separate sisters, but the separating activity is not turned on until Cut2/Pds1 is cleaved and destroyed^{7–9}.

Further genetic screens identified proteins that bind to each other to form the cohesin complex, so named because sister chromosomes separate prematurely when any of its members are inactivated^{10–12}. Two subunits of this complex, Smc1 and Smc3, are members of a class of ATPases that can alter the three-dimensional path of DNA molecules in space and have been implicated in a wide range of large-scale chromosomal behaviours including chromosome condensation, sister chromosome linkage, recombination and dosage compensation (reviewed

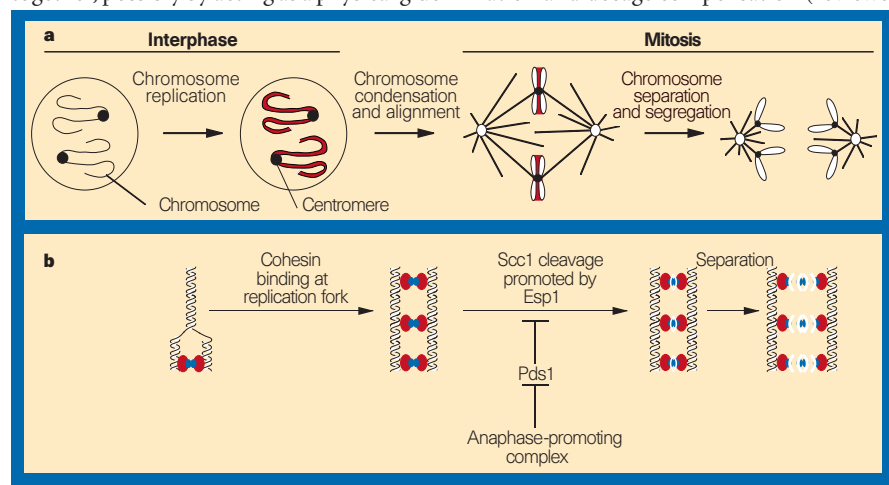


Figure 1 Chromosomal events during cell division. a, Essential steps. During interphase, DNA replication generates two sister chromosomes, which are linked together by a molecular 'glue' (shown in red). The chromosomes condense and line up on the mitotic spindle, then the two sisters separate and migrate to opposite poles of the cell. This is followed by cell cleavage, yielding two genetically identical daughter cells. b, Molecular details. The different stages are aligned beneath the corresponding stages of the chromosome cycle. The cohesin complex (red) binds to the sister chromosomes during DNA replication. According to the results of Uhlmann *et al.*¹, the Scc1 subunit (blue) of the cohesin complex is cleaved at mitosis. This reaction is promoted by Esp1 which, in turn, is freed up to do this job through destruction of another protein called Pds1 by the anaphase-promoting complex. Cleavage of the cohesin complex allows the sister chromosomes to separate, and cell division to proceed.

in ref. 13). The other proteins in the cohesin complex lack any known enzymatic activity, but are required for the complex to form and to bind to DNA. One of these subunits, Scc1 (also called Mcd1), dissociates when the sister chromosomes separate^{10,11}, but only if Esp1 is active. These results suggest that the cohesin complex holds the sisters together, and that its removal from the chromosomes triggers sister separation⁹. The complex can establish this linkage only during DNA replication, possibly owing to activities that travel with the replication fork^{12,14}.

These studies set the stage for the work of Uhlmann and colleagues¹, who used a mixture of biochemistry and genetics to show that sister separation requires cleavage of Scc1. The biochemical analysis relied on extractology — the study of complex reactions in crude, but concentrated, cell extracts that can be manipulated by the addition or subtraction of individual proteins. When Uhlmann *et al.* added chromatin to such an extract, the Scc1 was cleaved into three fragments and it dissociated from the chromatin. The authors found that this reaction requires Esp1 and is inhibited by Pds1, reflecting the requirements for sister separation *in vivo*.

Is cleavage a cause or a consequence of sister separation? To answer this critical question, Uhlmann *et al.* mapped the cleavage sites, then made mutations that blocked them and reintroduced the mutant gene into cells. The results were dramatic. Although the mutant Scc1 protein established sister-chromosome linkage during DNA replication, it did not leave the chromosomes during mitosis, and the cells died because the linkage between the sisters was not dissolved. So, two different sorts of proteolysis are needed to initiate sister separation. The first is activation of the anaphase-promoting complex, which leads to the wholesale destruction of Pds1. This, in turn, frees Esp1 to introduce two surgical snips in Scc1, thereby destroying the cohesin complex (Fig. 1b).

Other studies indicate that changes in the cohesin subunits are responsible for the difference between chromosome segregation in mitosis and meiosis. In mitosis, the sisters separate their arms and centromeres (the specialized region that attaches them to the spindle) at the same time. But in meiosis, the arms separate in the first division whereas the centromeres separate in the second. In fission yeast, delayed separation of the centromeres can be induced by replacing components of the mitotic cohesion complex with variants that are made only during meiosis. This result indicates that a change in a single chromosomal protein may be enough to cause the altered pattern of chromosome segregation that is responsible for sexual reproduction (Y. Watanabe and P. Nurse, personal communication).

So, have we finally identified the physical glue that holds the sisters together? And is

Esp1 the protease that destroys it? The answer to both questions is a resounding 'maybe'. We cannot exclude the possibility that the cohesin complex regulates the stability of some other, more fundamental linkage. Support for this possibility comes from studies of frog egg extracts, in which the cohesin complex helps to set up sister linkage, but leaves the chromosomes as they enter mitosis — well before the sisters separate¹⁵. Moreover, sequence gazing suggests that Esp1 is unlikely to cleave Scc1, because it lacks homology to any known protease. The answers probably lie beyond the realms of genetics and extractology, but the questions should stimulate biochemists to roll up their sleeves and dissect the reactions that promote cell reproduction by abolishing chromosomal sisterhood. □

Andrew Murray is in the Department of Physiology, University of California, 513 Parnassus Avenue,

San Francisco, California 94143-0444, USA.

e-mail: amurray@socrates.ucsf.edu

1. Uhlmann, F., Lottspeich, F. & Nasmyth, K. *Nature* **400**, 37–42 (1999).
2. Holloway, S. L., Glotzer, M., King, R. W. & Murray, A. W. *Cell* **73**, 1393–1402 (1993).
3. Uzawa, S. & Yanagida, M. *J. Cell Sci.* **101**, 267–275 (1992).
4. Guacci, V., Hogan, E. & Koshland, D. *J. Cell Biol.* **125**, 517–530 (1994).
5. Straight, A. F., Belmont, A. S., Robinett, C. C. & Murray, A. W. *Curr. Biol.* **6**, 1599–1608 (1996).
6. Yamamoto, Y. A., Guacci, V. & Koshland, D. *J. Cell Biol.* **133**, 99–110 (1996).
7. Funabiki, H. *et al.* *Nature* **381**, 438–441 (1996).
8. Kumada, K. *et al.* *Curr. Biol.* **8**, 633–641 (1998).
9. Ciosk, R. *et al.* *Cell* **93**, 1067–1076 (1998).
10. Guacci, V., Koshland, D. & Strunnikov, A. *Cell* **91**, 47–57 (1997).
11. Michaelis, C., Ciosk, R. & Nasmyth, K. *Cell* **91**, 35–45 (1997).
12. Toth, A. *et al.* *Genes Dev.* **13**, 320–333 (1999).
13. Hirano, T. *Genes Dev.* **13**, 11–19 (1999).
14. Uhlmann, F. & Nasmyth, K. *Curr. Biol.* **8**, 1095–1101 (1998).
15. Losada, A., Hirano, M. & Hirano, T. *Genes Dev.* **12**, 1986–1997 (1998).

Particle physics

And you're glue

Frank Wilczek

It's a widely believed half-truth that protons and neutrons are made out of quarks. Actually, physicists are increasingly discovering that it's considerably less than half the truth. The modern theory of the strong force, which binds quarks inside protons and neutrons, and these particles in turn to make atomic nuclei, is quantum chromodynamics (QCD). The other ingredients of QCD, the colour gluons, were once conceived as mere paste that somehow links together more substantial stuff (their name reflects this). No longer. On closer inspection, the quarks appear as the showier, but gluons as the weightier and more dynamic, constituents of matter. Definitive images¹ from a microscope capable of looking inside protons, the HERA accelerator in Hamburg, Germany, reveal as well that there is more to gluons than meets the eye.

To understand these evolving views, you must consider how one goes about looking inside a proton, to 'see' what it is made of. An ordinary microscope, using ordinary light, is woefully inadequate, because the wavelength of light is about one billion times larger than the size of the proton. Even fancy electron or scanning tunnelling microscopes can barely resolve single atoms, and fall far short of seeing the nucleus inside. The right tool for the job is a high-energy accelerator. They produce virtual photons of very short wavelength (and lifetime), that can be used to take snapshots of the proton's interior (Box 1, overleaf).

There's a catch, however, to this seemingly straightforward procedure. You get to see only what the virtual photon allows you to see. And because the photons couple only to electrically charged particles, constituents of

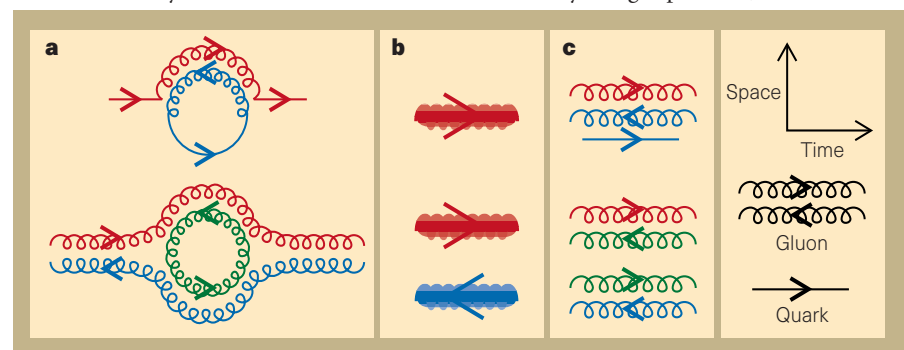


Figure 1 What once appeared as fuzzy quarks or gluons are clearly imaged in a faster exposure, which reveals additional gluons. a, The flow of colour charge around a quark (top) or gluon (bottom) in time. Quarks carry a single unit of colour charge: red, blue or green. Gluons carry both positive and negative units of colour charge. b, Average views of the same quark and gluon with coarse time resolution. c, In each case, a sharper-resolution view of the central time interval reveals the existence of an additional gluon.



100 YEARS AGO

A few interesting facts with regard to the kea, or sheep-eating parrot, of New Zealand are related in the July number of *Leisure Hour* by Dr. F. Truby King. The intense curiosity of these birds is stated to be sufficient to account for the habit of eating sheep acquired by them. Dr. King thinks it is probably a mistake to suppose that the kea designedly makes at once for the kidney fat of the sheep upon which it has pounced. It eats into various parts of the body, though perhaps more often into the region of the kidney, as it is there that the kea gets the firmest stand on the back of the running sheep. This view is strengthened by the fact that the bird prefers double-fleeced sheep – that is, such as have remained a whole season unshorn, on which it obtains a firmer grip.

From *Nature* 29 June 1899.

50 YEARS AGO

According to Darwin the long neck of the giraffe is the result of natural selection acting through the animal's tree-feeding habit. He wrote: "the individuals which were the highest browsers and were able during dearths to reach even an inch or two above the others will often have been preserved." There are serious objections to this argument. (1) During dearths ... the recurrent wastage of young giraffes would threaten the species with extinction. (2) Under such extreme conditions of dearth the grass-eating African ungulates would also have been so short of food that it is difficult to see why more of them did not develop the leaf-eating habit and the excessively long neck. (3) Bull giraffes tend to be several inches taller than the cows, so that during each dearth males would be naturally selected at the expense of the females – another factor likely to lead to rapid extinction. An alternative theory free from these criticisms has occurred to me. ... The giraffe is actively preyed upon by lions and leopards. It is reasonable, therefore, to explain the excessive length of its forelegs as the effect of natural selection acting continually through the hunter-hunted relationship. ... That the neck has elongated to a degree only just sufficient to keep pace with the increasing length of the legs is suggested by the fact that the giraffe has to splay its forelimbs awkwardly to drink.

From *Nature* 2 July 1949.

Box 1: Designer photons

The basic principle of the powerful microscopes used to examine proton interiors is extremely simple. Electrons (or positrons) are accelerated to high energy, and made to collide with protons (or other atomic nuclei). The main interactions of electrons are through the well-understood processes of electrodynamics. In quantum electrodynamics (QED), its modern version, these interactions are pictured and calculated as the exchange of real or virtual photons.

Real photons are completely characterized by their energy. Their momentum is simply equal to their energy (divided by the speed of light). For virtual photons, which exist only for a limited time, the energy and momentum are independent variables. In either case, real or virtual, the wavelength is inversely proportional

to the momentum.

When an electron at SLAC interacts with a proton it generally gets deflected, so that its energy and momentum change. The difference is carried off by a virtual photon. If the momentum of this virtual photon is large, it will have a short wavelength and provide excellent spatial resolution. If, in addition, it is a 'highly virtual' photon – that is, if there is a big mismatch between its energy (divided by the speed of light) and momentum, then it has a very short lifetime and provides excellent time resolution. By analysing how frequently virtual photons of different kinds get absorbed – that is, by noting how frequently you observe different electron deflections – you can effectively take snapshots, with adjustable spatial and time resolutions, of the charge distribution inside protons.

F.W.

the proton that are electrically neutral do not show up in these snapshots. (Biologists face a similar problem, in that they are often interested in materials that are basically transparent. They use various tricks, such as staining or attaching fluorescent molecules, to work around this difficulty.) What you see in an accelerator is quarks². Although these unusual particles are notoriously 'well confined', that is, impossible to isolate, they show up quite clearly in these short-time-exposure snapshots. You can even measure their peculiar fractional electric charges, their spin and the energy and momentum they carry.

When this is done, and the results analysed, physicists discover that they have a missing-mass problem on their hands,

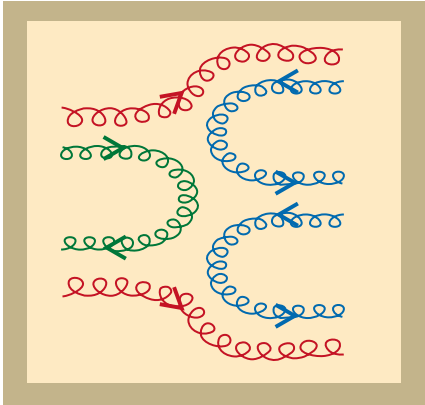


Figure 2 Recombination of gluons. In Fig. 1, a single smudgy gluon was resolved into two sharp ones. If gluon seas were always dilute, then two separate smudgy gluons should be resolved into four sharp ones. But when gluons are more densely packed (as shown), two could be resolved into three, suggesting recombination has taken place.

rather similar to the one faced by astronomers. The total energy in all the quarks does not add up to the energy in the proton. It is less than half. In the early days of this subject, Richard Feynman (who first carried out this kind of analysis) greatly annoyed his colleague Murray Gell Mann (a paladin of quarks) by calling the constituents of protons 'partons', and refusing to acknowledge that partons are quarks. As it turns out, each rival had part of the truth. Some of the partons are quarks, but many are gluons. The quarks carry all the electric charge, but most of the energy is in the colour gluons.

So what's new? The latest and greatest accelerators allow us to take snapshots with ever shorter time exposures. We can thereby resolve in ever finer detail the dynamical processes going on inside protons. When this is done, a remarkable result emerges. In the shorter-exposure pictures, one finds that the balance of energy tilts more and more toward gluons. What previously appeared as a blurry quark is on sharper examination revealed as part quark, part gluon. Nor is the concept of gluons free from challenge. They, too, on sharper examination, are revealed to be composites of several lower-energy gluons (Fig. 1). The closer we look, the more a proton (or neutron) appears as a bundle of soft glue¹.

This remarkable phenomenon – the gluonization of the proton – was predicted theoretically³. Indeed, it reflects quite directly one of the most profound features of the fundamental theory of quarks and gluons. According to QCD, the very powerful, complicated forces that confine quarks, preventing them from drifting away over long periods of time, are due to the accumulated effect of a surrounding cloud of short-lived virtual

gluons, each of which individually interacts weakly and simply. Now we've got pictures to prove it.

Predicted as a possibility, but not yet observed, is the inverse process, where gluons become so densely packed that they start recombining. This happens when two nearby fuzzy gluons are resolved, on closer inspection, into three (rather than the expected four — two from each) sharp gluons (Fig. 2). It can even turn out that false images are dissolved away: what appears at first as three fuzzy-looking gluons can turn out to be the image of two less fuzzy ones that partially overlap. A fascinating equilibrium of splitting and recombination should eventually result⁴. It is a considerable theoretical challenge to calculate this equilibrium in detail — how a proton eventually appears, as you peer in closer and closer.

In an accelerator, rapidly approaching nuclei appear to one another as pancakes, because of relativistic contraction, and thereby produce the densest gluon seas⁵. Their use, now an active frontier of experimental physics, will open the recombination regime to experimental investigation. The

accumulation of glue inside the proton is not only of interest for its own sake. It will also be important for the interpretation of future experiments using ultra-high-energy protons as an exploratory tool, particularly searches for the Higgs particle and the particles predicted by supersymmetry at the CERN Large Hadron Collider. Because many of these particles are produced primarily by the fusion of gluons found within colliding protons, their production rate is controlled by the gluons' densities⁶.

In any duel of insults, the old saying, "I'm rubber, and you're glue, so whatever you call me bounces back and sticks to you", is a feeble last resort. Well, you are, you know. □

Frank Wilczek is at the Institute for Advanced Study, School of Natural Sciences, Olden Lane, Princeton, New Jersey 08540, USA.

e-mail: wilczek@sns.ias.edu

1. Abramowicz, H. & Caldwell, A. *Rev. Mod. Phys.* (submitted); <http://xxx.lanl.gov/abs/hep-ex/9903037>
2. Roberts, R. *The Structure of the Proton* (Cambridge Univ. Press, 1990).
3. De Rujula, A. *et al. Phys. Rev.* **10**, 1649–1652 (1974).
4. Gribov, L., Levin, E. & Ryskin, M. *Phys. Reps* **100**, 1–146 (1983).
5. McLerran, L. <http://xxx.lanl.gov/abs/hep-ph/9903536>
6. Kilian, W. & Zerwas, P. <http://xxx.lanl.gov/abs/hep-ph/9809486>

Palaeobiology

A refugium for relicts

Zhexi Luo

The transition between the Late Jurassic and Early Cretaceous, 157–100 million years (Myr) ago, is a defining time in the history of terrestrial biodiversity. It witnessed the descent of birds from two-legged, meat-eating dinosaurs^{1,2}, and their early diversification^{3–5}; the rapid evolution of non-bird dinosaurs; diversification of the major mammalian groups^{6,7}; the origins of flowering plants (angiosperms)⁸; and diversification of the nectar-feeding flies — the finest early examples of plant–insect co-evolution^{9,10}. These emerging lineages came to dominate the world's terrestrial biotas, and some are still thriving today.

A window with a grand view of this evolutionary spectacle is the Yixian Formation in China's Liaoning Province, one of the richest and most important sources of fossils from the Mesozoic era (245–65 Myr ago; see map on page 59). On page 58 of this issue¹¹, Carl Swisher, Yuan-qing Wang and their colleagues show that the Yixian beds are 124 Myr old, placing them firmly within the Early Cretaceous. Such precisely dated rich fossil assemblages are rare in the Early Cretaceous¹², making the new work a welcome step towards a better temporal calibration of the Mesozoic terrestrial biotas.

A cornucopia of beautiful fossils has been unearthed, in unprecedented quantities, from the Yixian Formation. The discovery of such exceptionally complete fossils has filled

the gaps in our knowledge about lineages that were previously represented only by sparse, fragmentary fossils. Although the Yixian beds have yielded fossil fish, plants and arthropods since the 1920s, most of these species are restricted to eastern and central Asia. Likewise, the abundant dinosaurs, such as the horned psittacosaur and enigmatic therizinosaurs found in the Yixian Formation, were also endemic to Asia during the earliest Cretaceous¹³. Without

fossils from outside Asia, it has been difficult to correlate the Yixian biota with the worldwide geological timescale.

Ideally, to date fossil beds we need 'index' fossils that are abundant, broadly distributed and have rapid evolution. Age correlation of terrestrial biotas — as in the case of the Yixian — can be more difficult, because the preservation of terrestrial fossils is rarer, less complete and more uneven geographically than in marine environments. So, for decades, divergent opinions have waxed and waned about the age of the Yixian biota, which was suggested to be either Late Jurassic or Early Cretaceous — or anywhere in between.

The controversy has implications for both temporal and geographical patterns in the evolution of Mesozoic plants and animals, as many fossils from the Yixian beds occupy critical positions on their respective family trees. For instance, *Sinosauropteryx*¹ is the most primitive coelurosaurian dinosaur with downy feathers. *Protarchaeopteryx* and *Caudipteryx*² are basal in the maniraptoran dinosaur lineage that also includes *Archaeopteryx* and other birds, suggesting that true feathers evolved before the origin of avian flight. Other examples include the symmetrodont mammal *Zhangheotherium*⁶, which is basal to the diverse spalacotheriid tree¹⁴, and *Jeholodens*, a close relative of the triconodontids⁷ that were found around the world in the Late Jurassic and Early Cretaceous. Then there are *Confuciusornis* and *Changchengornis*, the primitive beaked birds capable of powered flight, which are relatives of the enantiornithine birds that dominated the global avian faunas in the Cretaceous^{3,15}.

Similar examples abound in plants and invertebrates. The primitive angiosperm *Archaeofructus*, for instance, has a typical angiosperm carpel (the female reproductive structure) but no flower petals⁸. It tells us much about the early anatomical evolution

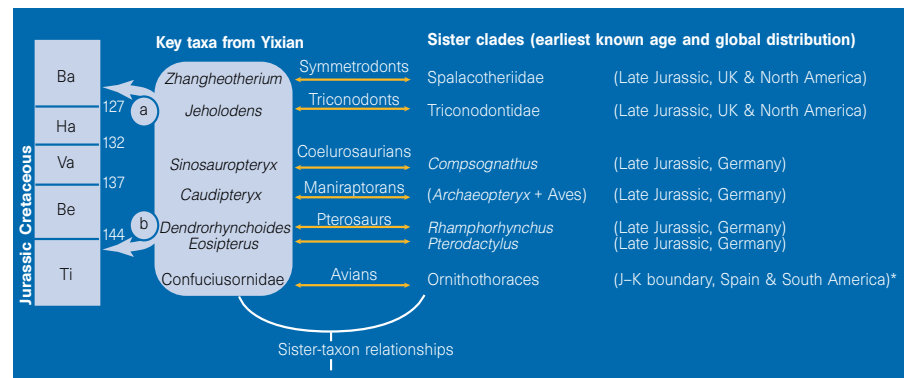


Figure 1 Correlation of taxa from the Yixian Formation with their close relatives from the Lower Jurassic or earliest Cretaceous in other continents. Several endemic taxa, such as *Psittacosaurus*, therizinosaurs and *Liaoningornis*, are not considered here, because they were restricted to northern China and Central Asia before the end of the Early Cretaceous, and cannot be correlated with the taxa on other continents. a, Correlation of the Yixian Formation, according to the results of Swisher *et al.*¹¹. b, Correlation of vertebrates from Yixian with their respective sister taxa from the Late Jurassic on other continents. Epochs are: Ba, Barremian; Ha, Hauterivian; Va, Valanginian; Be, Berriasian; Ti, Tithonian.

of flowering plants. Angiosperm plants and the flies that pollinate them are thought to have first diversified in the temperate climate of East and Central Asia, before spreading elsewhere^{8,9}. This would be consistent with a Late Jurassic age for the Yixian Formation, and could mean that northeastern China was a source for the emergence and early diversification of many evolutionary lineages that dominated the world's biota in the Cretaceous and the Cenozoic (65 Myr to present).

In harmony with this view, several vertebrates from the Yixian Formation have been shown to be closely clustered with European and North American relatives from the Late Jurassic (Fig. 1). For instance, the flying reptiles (pterosaurs) from Yixian closely resemble those of the Jurassic Solnhofen limestone in Germany¹⁶. Likewise, the Yixian Formation's *Sinosauropteryx* and Solnhofen's *Compsognathus* are very similar^{1,16}. The two new mammals from Yixian, *Zhangheotherium*^{6,14} and *Jeholodens*⁷, are both related to groups common in the Late Jurassic of Europe and North America. Such a parsimonious correlation by vertebrate sister taxa¹⁷ on an intercontinental scale tends to suggest a Late Jurassic age for the Yixian fossil assemblage.

But Swisher *et al.*¹¹ have now found that the sediments that yielded the feathered dinosaurs and mammals in the Yixian Formation date back to the Early Cretaceous. They have based this figure on measurements taken from sanidine, a mineral that is known to be reliable for determining the ⁴⁰Ar/³⁹Ar isotope age. The authors collected samples from the volcanic tuffs that are a part of the original sediments that bear the fossils — so, the association of the dated samples and the fossils is unambiguous.

All of this means that the primitive organisms from the Yixian Formation are best interpreted as long-lived relicts, extending from the Late Jurassic into the Early Cretaceous¹¹. A corollary is that the western Liaoning province and its neighbouring areas are a refugium — an isolated area in which a population survived much longer than elsewhere — for these relict lineages from a bygone era. In other words, *Sinosauropteryx*, the primitive pterosaurs and pre-flower angiosperm plants should be considered as 'living fossils' of their time. They survived into the Early Cretaceous only in Yixian, having previously been widely distributed around the world in the Late Jurassic. This agrees with the observation that non-theropod dinosaur faunas of eastern and central Asia were isolated from the rest of the world during the Late Jurassic and earliest Cretaceous¹³.

Primitive fossil taxa may not always be older than their more derived relatives. Terrestrial fossils are often incompletely preserved and, of those that are, only a fraction are ever discovered. Where the primitive

'stage of evolution' for a taxon does not match up with its apparently young age (compared with its more derived relatives), it is reasonable to infer that this taxon had a 'ghost lineage', or an earlier history that must have existed but is not yet documented by the fossil record¹⁸. So, an Early Cretaceous age for the Yixian Formation is not inconsistent with its seemingly archaic biota, in view of its restricted geographical distribution in China and Central Asia.

Had Swisher and colleagues come up with a Late Jurassic age, the Yixian Formation could have been viewed as a paradise for the origination and early diversification of many groups that are still thriving today. As it is, though, the Yixian beds are more likely to have been a geographic refugium for many relict lineages of the Jurassic. □

Zhexi Luo is in the Section of Vertebrate Paleontology, Carnegie Museum of Natural History, Pittsburgh, Pennsylvania 15213-4080, USA.
e-mail: luoz@clpgh.org

1. Chen, P.-J., Dong, Z.-M. & Zhen, S.-N. *Nature* **391**, 147–152 (1998).
2. Ji, Q., Currie, P. J., Norell, M. A. & Ji, S.-A. *Nature* **393**, 753–761 (1998).
3. Hou, L., Zhou, Z., Martin, L. D. & Feduccia, A. *Nature* **377**, 616–618 (1995).
4. Hou, L., Martin, L. D., Zhou, Z. & Feduccia, A. *Science* **274**, 1164–1165 (1996).
5. Ji, Q., Chiappe, L. M. & Ji, S.-A. *J. Vert. Paleontol.* **19**, 1–7 (1999).
6. Hu, Y., Wang, Y., Luo, Z. & Li, C. *Nature* **390**, 137–142 (1997).
7. Ji, Q., Luo, Z. & Ji, S. *Nature* **398**, 326–330 (1999).
8. Sun, G., Dilcher, D. L., Zhen, S.-L. & Zhou, Z.-K. *Science* **274**, 1164–1167 (1998).
9. Ren, D. *Science* **280**, 85–88 (1998).
10. Labandeira, C. C. *Science* **280**, 57–58 (1998).
11. Swisher, C. C. III, Wang, Y.-Q., Wang, X.-L., Xu, X. & Wang, Y. *Nature* **400**, 58–61 (1999).
12. Gradstein, F. M. *et al.* in *Geochronology, Time Scales and Global Stratigraphic Correlation* (eds Berggren, W. A., Kent, D. V., Aubry, M. P. & Hardenbol, J.) 95–126 (SEPM, Spec. Publ. No. 54, 1995).
13. Russell, D. A. *Can. J. Earth Sci.* **30**, 2002–2012 (1993).
14. Cifelli, R. L. & Madsen, S. K. *Geodiversitas* (in the press).
15. Chiappe, L. M. *Nature* **378**, 349–355 (1995).
16. Ji, S.-A., Ji, Q. & Padian, K. *Nature* **398**, 573–574 (1999).
17. Padian, K., Lindberg, D. R. & Polly, D. P. *Annu. Rev. Earth Planet. Sci.* **22**, 63–91 (1995).
18. Norell, M. A. *Am. J. Sci.* **293A**, 407–417 (1993).

Malaria

Tying the conductor's arms

Michael F. Good

To survive their hosts' immunological defences, microorganisms have developed a plethora of strategies that suppress or subvert the immune system¹. The most successful organisms manage to fine-tune their hosts' immune responses — they suppress these responses enough to allow their own survival, but leave some immune surveillance to prevent uncontrolled growth (which would kill the host). In this way, the microorganism can be passed to another

host, so continuing the species. Take, for example, the malaria parasite *Plasmodium falciparum*, which inhabits red blood cells and is passed from host to host by mosquitoes. For such transmission to occur, there must be a considerable amount of the parasite in the blood, meaning that the immune response needs to be severely checked^{2,3}. And on page 73 of this issue, Urban *et al.*⁴ describe a newly discovered way in which *P. falciparum* does this.

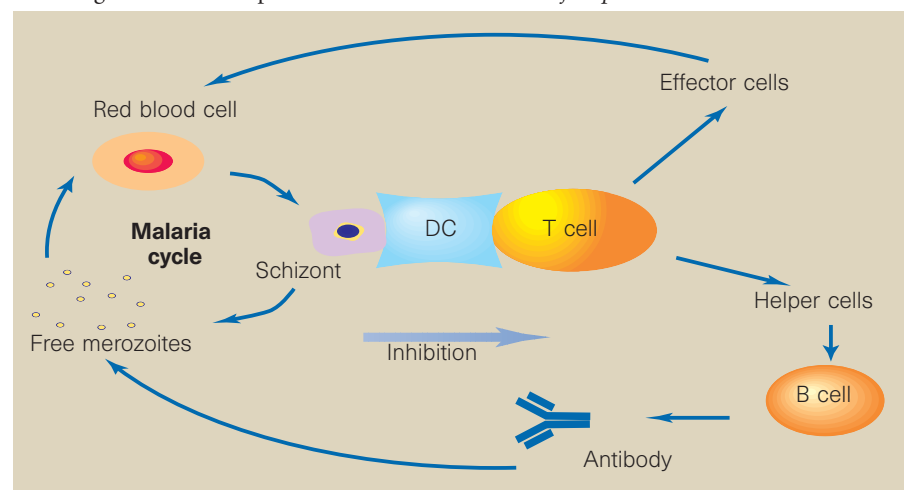


Figure 1 Effect of malaria-infected red blood cells on the host immune response. The malaria parasite *Plasmodium falciparum* invades red blood cells and, via spores called schizonts and through the formation of mitotic products called merozoites, invades other host cells. Urban *et al.*⁴ have now shown that *P. falciparum* also suppresses the ability of dendritic cells (DC) to initiate immune responses. This means that neither the effector T cells (which fight the parasite via cytokines and other inflammatory molecules) nor the helper T cells (which fight the parasite via binding of antibodies to merozoites and infected red blood cells) are activated. The immune responses shown are parasite specific, but malaria infection can also suppress immune responses to other organisms.

People with malaria have a reduced immune response not only to the malaria parasite, but also to other organisms and antigens (foreign proteins), including vaccines^{5,6}, possibly because antigens of the malaria parasites are very similar to those of other microorganisms⁷. To impede the immune response strongly, the parasite must target it at an early stage. Several studies have hinted that immune responses in malaria might be impaired because fewer macrophages are generated in the bone marrow, and because these smaller populations process antigen ineffectively^{8,9}. Macrophages, along with other 'antigen-presenting cells', can start or re-boot the immune response by processing and then presenting fragments of antigens to the T cells, of which there are two types. The 'effector' T cells kill the microbe directly, whereas the 'helper' T cells enable B cells to make antibody against it (Fig. 1).

The most potent antigen-presenting cells are the so-called dendritic cells¹⁰. Found in all tissues, dendritic cells have the unique ability not only to assist in boosting or amplifying an immune response on subsequent exposure to an organism, but also to initiate a response against a new microbe. Clearly, any suppression of these cells would result in something akin to an orchestra attempting to learn a new symphony when their conductor's arms were tied. Yet this is what Urban and colleagues⁴ now show that *P. falciparum* does.

The authors have found that *P. falciparum* prevents the maturation of dendritic cells. As a result, functional activity of the dendritic cells is impeded — Urban *et al.* measured dramatic reductions in T-cell proliferative responsiveness to the malaria parasite and other antigens. The authors show that binding of infected intact red blood cells to the surface of the dendritic cells is critical for impairing the activity of dendritic cells. We already know that the ability of malaria-infected red cells to stick to the endothelium of blood vessels in the brain and other tissues (such as the placenta¹¹) is related to disease pathogenesis. Indeed, some of the parasite proteins involved in adhesion have been defined, including an antigen known as PfEMP1. Molecules that bind to this antigen in endothelial cells have also been identified, and include CD36 and thrombospondin.

Urban and colleagues suggest that the PfEMP1 antigen may be responsible for binding to dendritic cells. Two parasite lines that adhere to CD36 and thrombospondin also stick to dendritic cells and inhibit their maturation. But a line that does not bind to CD36 and thrombospondin does not bind to dendritic cells and does not affect their function. Their data do not, however, exclude the possibility that other antigens from the parasite bind to other ligands on the dendritic

cells. Blocking studies with antibodies may identify the critical interactions. Moreover, although vascular adhesion is a well-known property of *P. falciparum*-infected red blood cells, it has not been seen with another strain of malaria parasite, *P. vivax* — yet immunosuppression has been described in both types of infection^{2,3}.

Urban and co-workers think that adhesion of infected red blood cells to dendritic cells is not a chance event. Instead, they say, it results from a selective advantage to the parasite, such that it is not destroyed by the immune response. But because the immune response is affected at such a critical, central place, it is not surprising that it has been difficult to fine-tune immune suppression to the combined advantage of both the parasite and the host. It is estimated that *P. falciparum* kills between 1.5 and 2.7 million children per year¹². Without drug intervention, of course, the death rate would be considerably higher — to neither the host's nor the parasite's advantage.

Although PfEMP1 is a variant antigen (that is, several different forms may be expressed over the course of a single infection), we could, perhaps, find conserved domains that would be the basis of a vaccine. But antibodies induced by specific vaccination should block adherence of the parasite, not only to blood vessels but also to dendritic cells. The result may be potent anti-parasite immune responses that clear the infection.

Clearly, other molecular pathways may be involved in malaria-induced immunosuppression, apart from those suggested by Urban and colleagues. Nevertheless, the authors do give us an insight into the immunoregulatory pathway(s) that the malaria parasite has adopted to survive. From these findings we can design strategies to stimulate immune responses to a variety of infectious organisms, or to dampen immune responses to self-antigens in the context of autoimmunity. □

Michael F. Good is at the Cooperative Research Centre for Vaccine Technology, Queensland Institute of Medical Research, PO Royal Brisbane Hospital 4029, Australia.

e-mail: michaelG@qimr.edu.au

1. Bloom, B. R. *Nature* **279**, 21–26 (1979).
2. Ho, M. *et al.* *J. Infect. Dis.* **153**, 763–771 (1986).
3. Goonewardene, R. *et al.* *Eur. J. Immunol.* **20**, 1387–1391 (1990).
4. Urban, B. C. *et al.* *Nature* **400**, 73–77 (1999).
5. Viens, P. *et al.* *Am. J. Trop. Med. Hyg.* **23**, 846–849 (1974).
6. Williamson, W. A. & Greenwood, B. M. *Lancet* **1**, 1328–1329 (1978).
7. Currier, J. *et al.* *Int. Immunol.* **4**, 985–994 (1992).
8. Silverman, P. H. *et al.* *Blood* **69**, 408–413 (1987).
9. Lechuk, R. & Playfair, J. H. *Clin. Exp. Immunol.* **42**, 428–435 (1980).
10. Cella, M. *et al.* *Curr. Opin. Immunol.* **9**, 10–16 (1997).
11. Fried, M. *et al.* *Nature* **395**, 851–852 (1998).
12. World Health Organization *The World Health Report 1998* 45 (WHO, Geneva, 1998).

Daedalus

Virtual overpopulation

Animal reproduction is a strategic process. Creatures such as fish, whose eggs are metabolically cheap, can spawn them in thousands in the hope that a few will survive. Those with costlier offspring must be more prudent. Birds, for example, have to build a nest, lay eggs, incubate them and feed the hatched chicks for some time. If the parents produce too many eggs, they will be unable to find enough food to support themselves and all the chicks over that time. Some chicks will die, wasting scarce resources. Conversely, too few eggs represent a lost reproductive opportunity.

The birds can, presumably, assess the food resources of their environment fairly well. But how do they judge the number of rivals competing for those resources? One solution is 'mustering'. Starlings, ducks, and many other species, gather together in the breeding season into vast flocks. Each bird makes the most conspicuous display and loudest noise it can manage, to show its rivals the reproductive competition they must face. All the birds gain some sense of their population density, and can judge the most cost-effective number of eggs to lay.

The mustering of urban birds, especially starlings, is a great nuisance in many big cities. Numerous attempts to poison or discourage these pests have failed. Daedalus now plans to subvert their mustering. He will place sets of huge parallel mirrors across the most popular mustering sites. The birds will see their numbers reflected repeatedly to infinity, and will decide that they are vastly overpopulated. Concave mirrors could magnify the more distant images, making them ever larger and more intimidating. The cowed birds will therefore lay very few eggs, and the urban nuisance will steadily abate.

The strategy can clearly be elaborated. Amplifiers (with time-delay to avoid feedback howl) could exaggerate the birds' cries. The surveillance cameras in many cities could portray the birds to themselves on vast numbers of TV screens around their sites. But Daedalus now wonders if cities themselves represent an instinctive human mustering. The steady flood of people into huge, crowded, noisy cities, especially in the developing world, may subconsciously impress all their inhabitants with the intensity of the reproductive competition. With luck, the resulting instinctive response will defuse the population explosion without anyone having to make hard policy decisions.

David Jones

The inventor of modern science

James Bradley laid the foundations of modern science in his aunt's attic. His impressively precise astronomical measurements gave birth to experimental physics as we know it. The first in our series of millennium essays.

Freeman J. Dyson

Modern science began in 1729, when it became based on measurements of high precision. This makes James Bradley, who did the first high-precision measurements, the inventor of modern science. He was the first to understand that accurate measurement requires meticulous monitoring and control of possible sources of error. He was the first to record temperature and barometric pressure whenever he made an observation.

In 1729, Newton had been dead for two years. Only 120 years had passed since Galileo's little telescope had first been pointed at Jupiter and its satellites. Bradley had the good luck to have a widowed aunt who owned the house where she lived in Wanstead, east London. The house had a high roof, and the aunt allowed him to cut a hole in her roof and use the upper part of the house as an observatory. Her husband had also been a devotee of astronomy, so she was used to this sort of madness. Under the hole, Bradley suspended a slender three-inch refracting telescope, with a brass circular arc along which the eyepiece of the telescope could be moved with a micrometer screw.

The eyepiece was focused simultaneously

on a star overhead and on a fine crosswire. Bradley could move the star exactly onto the crosswire with the micrometer screw, and then read off the angle on the brass arc. At one place on the arc there was a notch leaving space for a vertical plumb-line made of wire with one hundredth of an inch diameter. We must assume that he had his aunt well trained so that she did not slam doors while observations were in progress.

Using his plumb-line and his micrometer screw, Bradley was able to measure the angle between a star and the local vertical with six-figure accuracy, with an error of one or two parts per million. He was the first person to measure anything with six-figure accuracy. He measured angles a hundred times more accurately than the astronomers of Newton's time. That was how he discovered aberration.

Aberration is the displacement of the image of a star in the sky due to the speed of the Earth in its orbit around the Sun. Bradley found that all the stars moved in elliptical paths in the sky, coming back to their original

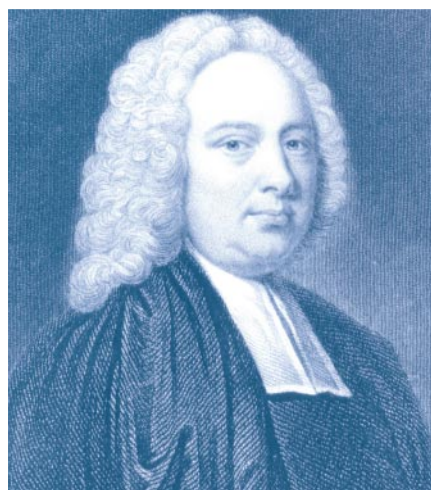
positions after a year. His discovery was acclaimed, not only by astronomers but by the educated public all over Europe, as the first direct demonstration that the Copernican view of the Universe was right. The elliptical motions of the stars in the sky are visual proof that the Earth is moving around the Sun. If the Earth were stationary with the Sun moving around it, aberration could not occur. The displacement in angle is equal to the Earth's velocity divided by the velocity of light. The magnitude of the observed displacements, when compared with the known orbital speed of the Earth, gave the first accurate determination of the velocity of light. Bradley's value for the velocity of light is within one per cent of the modern value.

Bradley did not have students in the modern style, but he had assistants, young men whom he trained in precision measurement and calculation. One of his assistants was Charles Mason, who afterwards achieved immortality as surveyor of the Mason–Dixon line separating Maryland from Pennsylvania. The line ran for 233 miles along the parallel of latitude $39^{\circ} 43'$, across land claimed by greedy landowners on both sides. It was important to stake out the line on the ground with astronomical precision. It was important that Mason knew about the disturbing effects of aberration on measurements of latitude. If aberration had not been taken into account, the line might have deviated from its proper position by as much as half a mile north or south, and the American Civil War might have started 100 years earlier than it did.

The influence of Bradley extended far beyond astronomy. The revolution that he started by insisting on six-figure accuracy spread to other countries and to other sciences. A succession of great mathematicians, from Laplace to Gauss and Poincaré, created the new science of analytical dynamics to impose a coherent intellectual order on the precisely observed celestial motions. One hundred and fifty years after Bradley, Albert Michelson and Edward Morley attempted to measure the second-order aberration of light caused by the Earth's motion through the ether, and by their negative result led the way to the theory of relativity.

Finally, after 200 years, the tradition of six-figure accuracy begun by Bradley led to the modern flowering of experimental physics, with Isidor Rabi's molecular beam apparatus probing the resonances of atoms as precisely as Bradley's micrometer probed the motions of the stars. □

Freeman J. Dyson is at the Institute for Advanced Study, Princeton, New Jersey 08540, USA.



Aberration becomes the norm: Bradley's careful measurements proved that the Earth orbits the Sun, and established the now traditional working methods of science.

CORBIS/BETTMANN

HUTTON GETTY

Genetic link to cervical tumours

Cervical cancer is strongly associated with infection by oncogenic types of human papilloma virus (HPV). But only a small fraction of those infected develop cancer, indicating that other factors contribute to the progression to cervical cancer. We have compared incidence of the disease in relatives of cases of cervical tumour and controls, and find a significant familial clustering among biological, but not adoptive, relatives. We find no difference in the risk to siblings who have a mother or father in common, so the clustering cannot be explained by vertical transmission of HPV from mother to child. These results provide epidemiological evidence of a genetic predisposition to cervical cancer.

We used the Swedish national registers to identify relatives of cases of cervical cancer (126,893 relatives of 71,533 cases) and randomly selected age-matched controls (334,961 relatives of 194,810 controls). From these data we estimated the familial relative risk (the risk for relatives of cases divided by risk for relatives of controls) of developing cervical dysplasia, cervical cancer *in situ* and invasive cancer. We searched the Swedish cancer register, which contains records of all cancers diagnosed in Sweden between 1958 and 1993, for all cases of these cervical tumours, according to the criteria of the seventh revision of the International Classification of Diseases. Most cases were diagnosed as carcinoma *in situ* (85%), with dysplasia and invasive cancer making up 10% and 5% of cases, respectively. The family register (Statistics, Sweden) is a population-based record of parents to individuals born after 1940.

For biological mothers to cases, the relative risk was 1.83 (95% confidence interval, 1.77–1.88), whereas for adoptive mothers the relative risk was not significantly different from unity (1.10, 0.76–1.54) (Fig. 1). The relative risk for biological full sisters to cases was 1.93 (1.85–2.01), whereas that for non-biological sisters to cases was not significantly different from unity (1.15, 0.82–1.57). These results indicate that there is a significantly increased risk to biological, but not to non-biological, first-degree relatives of women with cervical tumours (Fig. 1). Further support for a genetic contribution was obtained by the analysis of second-degree relatives. For half-sisters to cases, the relative risk was 1.45 (1.31–1.60). Because half-sisters share, on average, 25% of their genetic material, this estimate is close to that expected if the familial aggregation seen for first-degree relatives is caused mainly by genetic factors.

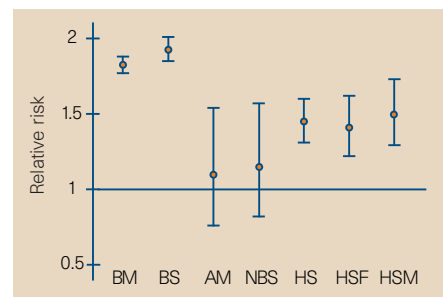
Vertical transmission of HPV from mother to child (during pregnancy or deliv-

Figure 1 Relative risk of cervical tumour (and 95% exact confidence interval) for relatives of cases. BM, biological mother; BS, biological sister; AM, adoptive mother; NBS, non-biological sister; HS, half-sister; HSF, half-sister with father in common; HSM, half-sister with mother in common. For each case with cervical tumour born between 1941 and 1977 in Sweden and diagnosed between 1958 and 1993, three age-matched controls were selected at random from the Swedish population. Cases or controls that lacked information about parents' identity

were not included. Controls diagnosed to have cervical cancer were excluded from the study. For cases with more than one type of diagnosis, we chose the most recent. Cohorts of distinct types of relatives to cases and controls were created and person-years at risk from 1958 to the date of a diagnosed cervical tumour, or until the end of follow-up (31 December 1993), were calculated. The five-year calendar time and age-specific incidences of cervical dysplasia, *in situ* carcinoma and invasive cancer for any given group of relatives to the controls were used to determine the expected number of affected individuals in the corresponding cohort of relatives to cases. The relative risk was then calculated by dividing the observed with the expected number of diagnoses in each cohort. Because the register allowed complete ascertainment, an affected sister of a case (proband) will appear once as a proband and at least once as a sister. To adjust the too-narrow confidence intervals arising from this multiple counting of sibships, both the observed and the expected number of cases were divided by two before calculation of the confidence intervals⁶. This correctly accounts for the multiple counting of sibships with two, but not more, affected sisters. However, only 3.6% of the multi-case sibships had more than two sisters affected, so this correction appears to be sufficiently conservative. The cohort of non-biological sisters consists of adopted sisters to not-adopted probands, and not-adopted sisters to adopted probands. Regional differences in incidence (from sources such as unequal diagnosis criteria among physicians at different hospitals, or differences in screening strategies) may bias the result and give rise to familial clustering. We therefore calculated prevalences for biological full sisters to proband cases, conditional on the county in which the proband cases were diagnosed. The prevalences ranged from 7.0% to 12.4% for different counties, whereas the average prevalence among sisters to controls was 4.9%. Thus, independent of the geographic region in which the proband cases were diagnosed, the prevalence of their sisters was statistically significant above the prevalence among sisters to healthy controls, showing that regional differences in incidence is not a critical source of bias.

ery) has been proposed as a risk factor for cervical cancer¹. If the familial clustering is caused by vertical transmission, the relative risk for half-sisters to affected individuals is expected to be greater when the mother, rather than the father, is the parent in common. However, the relative risk for half-sisters to cases having a father (1.41, 1.22–1.62) or mother (1.49, 1.29–1.73) in common are similar (Fig. 1). Thus the familial clustering cannot be explained by transmission of HPV from mother to child.

Familial cancers are reported to have an earlier age of onset than sporadic forms². By using the regression of age at diagnosis by year of birth, we found that familial cases with cervical cancer had a significantly lower age at diagnosis (intercept) than sporadic cases (no affected family member), both for daughters (0.6 years, $P < 0.05$) and sisters (0.6 years, $P < 0.05$). If relatives affect each others' tendency to go for medical examination, this could bias the analysis. However, no difference was found in the mean age at diagnosis between first-born and last-born sisters (data not shown). This difference in age at diagnosis between familial and sporadic cases lends further support for a genetic contribution to cervical tumours. The small



difference seen (less than one year) in the age at diagnosis between familial and sporadic cases would be expected if the genetic component acts mainly at the level of susceptibility to infection by the virus, affecting the rate of tumour progression to a much smaller extent.

Our results demonstrate a significant familial clustering of Swedish cases of cervical tumours which is more likely to result from genetic rather than environmental factors. Comparisons of mono- and dizygotic twins have also indicated a genetic contribution to cervical cancer *in situ*³. If the contribution of a shared environment to the clustering is substantial then the risk for siblings should be higher than for mother–daughter relations, as they share the environment during the same time and age period⁴. The difference in relative risk for the two types of relationships (1.93 for full siblings and 1.83 for mothers) is small, consistent with the familiarity being due mainly to genetic factors.

Our results therefore provide epidemiological evidence for a genetic predisposition to cervical cancer and its precursor forms, and support further searches for the causative loci. Indeed, the familial risk to

cervical tumours is of the same magnitude as that seen for other forms of cancer, such as prostate cancer³, for which a genetic contribution has long been recognized.

Patrik K. E. Magnusson*, Pär Sparén†, Ulf B. Gyllensten*

*Department of Genetics and Pathology,

Section for Medical Genetics, Box 589,

University of Uppsala,

S-751 23 Uppsala, Sweden

e-mail: ulf.gyllensten@medgen.uu.se

†Department of Medical Epidemiology,

Karolinska Institute, Box 281,

S-171 77 Stockholm, Sweden

1. Kaye, J. N. *et al.* *J. Gen. Virol.* **77**, 1139–1143 (1996).
2. Heimdal, K. *et al.* *Br. J. Cancer* **73**, 964–969 (1996).
3. Ahlbom, A. *et al.* *J. Natl Cancer Inst.* **89**, 287–293 (1997).
4. Carstensen, B., Soll-Johanning, H., Villadsen, E., Sondergaard, J. O. & Lyng, E. *Int. J. Cancer* **68**, 428–435 (1996).
5. Isaacs, S. D., Kiemeny, L. A., Baffoe-Bonnie, A., Beaty, T. H. & Walsh, P. C. *J. Natl Cancer Inst.* **87**, 991–996 (1995).
6. Goldgar, D. E., Easton, D. F., Cannon-Albright, L. A. & Skolnick, M. H. *J. Natl Cancer Inst.* **86**, 1600–1608 (1994).

Isochores result from mutation not selection

Mammalian genomes are made up of large regions of distinct base composition known as isochores¹, but it has remained uncertain whether this mosaic structure is caused by regional variation in mutational biases^{2–4} or by natural selection¹. To resolve this controversy, we compared the rate and pattern of nucleotide substitutions in two globin pseudogenes, $\Psi\alpha 1$ and $\Psi\eta$, which reside in isochores of very different base composition in the genomes of humans and other primates. Our findings indicate that isochores result from mutational biases, rather than from selection on individual nucleotides.

Vertebrate globin genes are divided into the α and β clusters, which arose through the duplication of an ancestral globin gene about 450 million years ago⁵. In mammals, the α and β clusters are situated within isochores of widely divergent base compositions on different chromosomes¹. Pseudogenes $\Psi\alpha 1$, of the GC-rich α cluster, and $\Psi\eta$, of the GC-poor β cluster, are present in all simian primates. $\Psi\alpha 1$ was silenced in the stem-simians⁶ and $\Psi\eta$ in the basal primates⁵. $\Psi\alpha 1$ and $\Psi\eta$ are therefore paralogues that have been evolving free of functional constraints within the simian lineage, but in distinctive GC-content backgrounds. Because nucleotide substitutions in regions under no selective constraints reflect the actual mutational process, paralogous pseudogenes are suitable for studying how patterns of mutations vary between isochores of divergent base composition⁷.

For each of the pseudogenes, we reconstructed the nucleotide substitutions among human (*Homo sapiens*), chimpanzee (*Pan troglodytes*), orang-utan (*Pongo pygmaeus*)

Table 1 Substitution patterns in $\Psi\alpha 1$ and $\Psi\eta$

	A	T	C	G
$\Psi\alpha 1$				
A	–	2.464	6.160	18.478
			($v=45.58$)	
T	4.927	–	14.783	6.160
C	3.936	10.231	–	2.362
		($u=33.84$)		
G	16.726	2.951	10.822	–
$\Psi\eta$				
A	–	5.458	4.777	15.012
			($v=34.21$)	
T	0.655	–	11.797	2.622
C	6.187	18.557	–	5.156
		($u=51.30$)		
G	22.535	4.026	3.220	–

Relative substitution frequencies in $\Psi\alpha 1$ (0.6 kilobases (kb), yielding 95 substitutions) and $\Psi\eta$ (1.5 kb, yielding 126 substitutions), where u is the relative substitution frequency from C/G to T/A, calculated as the sum of all relative substitution frequencies from C or G to T or A, and v is the substitution frequency from T/A to C/G. The v/u ratios of each of these pseudogenes ($v/u=1.35$ for $\Psi\alpha 1$ and 0.67 for $\Psi\eta$) show the opposing GC pressures in $\Psi\alpha 1$ and $\Psi\eta$. Nucleotide sequences of $\Psi\eta$ ¹¹ are available from GenBank (accession numbers U01317 (base pairs 45657–47124), K02542, M18038, J03818 and J03050). For $\Psi\alpha 1$, the sequence from human and the $A\alpha$ sequence from galago were obtained from GenBank (accession numbers J00153 (base pairs 2475–3075) and M29742), the sequences from chimpanzee and a partial sequence from cynomolgus monkey were from the literature¹², and the sequences from orang-utan, baboon and colobus, plus a fragment from cynomolgus monkey, were amplified by the polymerase chain reaction and sequenced. Multiple sequence alignments were performed with Clustal X¹³ and further refined by hand. Nucleotide substitutions were reconstructed by maximum parsimony⁸.

and one or more Old World monkeys: the cynomolgus monkey (*Macaca fascicularis*), olive baboon (*Papio anubis*) and black-and-white colobus (*Colobus guereza*) for $\Psi\alpha 1$, and rhesus monkey (*M. mulatta*) for $\Psi\eta$. The functional gene $A\alpha$ from the prosimian thick-tailed galago (*Galago crassicaudatus*), a $\Psi\alpha 1$ orthologue, and the pseudogene $\Psi\eta$ (*geoffroyi*) were used as outgroups. The ancestral states at the nodes of the phylogenies relating the sequences and the directionality of substitutions were inferred by maximum parsimony⁸.

The two pseudogenes undergo substitutional patterns of opposite direction: the ratio v/u is greater than 1 in $\Psi\alpha 1$ but less than 1 in $\Psi\eta$, where u is the relative substitution frequency from G/C to A/T, and v is the relative substitution frequency from A/T to G/C, expressed as percentages of the total number of substitutions and adjusted for the unequal representation of each nucleotide in the sequences^{3,9} (Table 1); the difference in v/u is statistically significant (log-odds ratio = 0.703 ± 0.304 , $P < 0.05$). In absolute frequencies, substitutions from A/T to G/C differ between the pseudogenes (26.85 in $\Psi\alpha 1$ and 15.09 in $\Psi\eta$, substitutions per thousand A/T; chi-squared test, $P < 0.01$), but substitutions from G/C to A/T do not (19.54 in $\Psi\alpha 1$ and 22.80 in $\Psi\eta$, per thousand G/C; chi-squared test, $P > 0.1$).

A mutational explanation for the existence of isochores requires that the bias in each region drive the GC content to its observed value. The equilibrium GC contents predicted by the substitutional patterns ($v/(u+v)$; ref. 3) are 57% in $\Psi\alpha 1$ and 40% in $\Psi\eta$, in agreement with the observed values of 59% and 43%. In each pseudogene, the actual numbers of substitutions

from G/C to A/T and from A/T to G/C do not differ statistically (38 and 37 in $\Psi\alpha 1$, 57 and 51 in $\Psi\eta$). Thus, the GC contents of $\Psi\alpha 1$ and $\Psi\eta$ are not changing with time, their equilibrium values being largely determined by the different mutational pressures affecting each pseudogene.

The lack of functional constraints on pseudogenes rules out any involvement of selection in shaping the substitutional pattern, with the possible exception of selection on DNA stability. Because G–C bonds are stronger than A–T bonds, it has been proposed that the emergence of GC-rich isochores in warm-blooded vertebrates is a selective response to an increased need for protecting DNA and RNA from heat degradation¹. If the GC-biased substitutional pattern in $\Psi\alpha 1$ has resulted from the selective elimination of mutations towards A/T, which would otherwise predominate in a GC-rich background under no directional GC pressure, observed rates of substitution would be reduced in the region. Furthermore, if GC-rich isochores were maintained by selection, while GC-poor isochores evolved neutrally, the underlying mutational bias over the entire genome would be towards A/T, and the effects of purifying selection in GC-rich regions would be stronger still. However, substitution rates are higher in GC-rich $\Psi\alpha 1$ than in GC-poor $\Psi\eta$ (4.38×10^{-9} against 2.62×10^{-9} substitutions per site per year, averaged over all shared species comparisons). Therefore, purifying selection is not affecting substantially the pattern of substitutions, which remains a good estimate of the mutational pattern.

Several mechanisms have been proposed to account for distinctive GC pressures among genomic regions, including

variation in replication time⁴, in chromatin state and transcriptional activity³, and in the machinery of DNA replication and repair¹⁰. Although natural selection could have played a role in creating such mechanisms, isochores result directly from mutational biases, rather than from selection acting on individual nucleotides.

M. Pilar Francino*, **Howard Ochman*†**

*Department of Biology,
University of Rochester,
Rochester, New York 14627, USA
†Department of Ecology and Evolutionary Biology,
University of Arizona,
Tucson, Arizona 85721, USA
e-mail: hochman@u.arizona.edu

1. Bernardi, G. *et al.* *Science* **228**, 953–958 (1985).
2. Filipiński, J. *FEBS Lett.* **217**, 184–186 (1987).
3. Sueoka, N. *Proc. Natl Acad. Sci. USA* **85**, 2653–2657 (1988).
4. Wolfe, K. H., Sharp, P. M. & Li, W.-H. *Nature* **337**, 283–285 (1989).
5. Goodman, M., Czelusniak, J., Koop, B. F., Tagle, D. & Slightom, J. L. *Cold Spring Harb. Symp. Quant. Biol.* **52**, 875–890 (1987).
6. Sawada, I. & Schmid, C. W. *J. Mol. Biol.* **192**, 693–709 (1986).
7. Casane, D., Boissinot, S., Chang, B. H.-J., Shimmin, L. C. & Li, W.-H. *J. Mol. Evol.* **45**, 216–226 (1997).
8. Maddison, W. P. & Maddison, D. R. *MacClade Version 3.0. Analysis of Phylogeny and Character Evolution* (Sinauer, Sunderland, MA, 1992).
9. Imanishi, T. & Gojobori, T. *J. Mol. Evol.* **35**, 196–204 (1992).
10. Holmquist, G. P. & Filipiński, J. *Trends Ecol. Evol.* **9**, 65–69 (1994).
11. Koop, B. F., Goodman, M., Xu, P., Chan, K. & Slightom, J. L. *Nature* **319**, 234–238 (1986).
12. Sawada, I. Thesis, Univ. California, Davis (1986).
13. Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F. & Higgins, D. G. *Nucleic Acids Res.* **25**, 4876–4882 (1997).

Periodicity in marine phosphorus burial rate

There have been arguments both for and against a periodicity of 26–33 million years (Myr) in terrestrial and extraterrestrial records^{1–5}. The best way to identify such periodicity is the analysis of geomarine evolutionary records⁶. We have analysed the marine sedimentary phosphorus burial rate (PBR), as fluctuations in this rate are strong indicators of the coupling of climate, continental weathering and ocean primary productivity⁷. We find a statistically significant harmonic component of 33 ± 3 Myr against the estimated robust background noise spectrum, supporting the idea that geomarine processes are cyclic.

The PBR records we analysed are based on published data from the Deep Sea Drilling Project (DSDP) and the Ocean Drilling Program (ODP), totalling 5,648 measurements⁷. A large number of sedimentary series were examined⁷, allowing the reconstruction of reliable PBR time series from different palaeoceanographic and sedimentological settings. The global nature of our data means we are unable to identify regional, physical or chemical sorting effects or ocean-to-ocean fractionations. Our chronological framework is based on

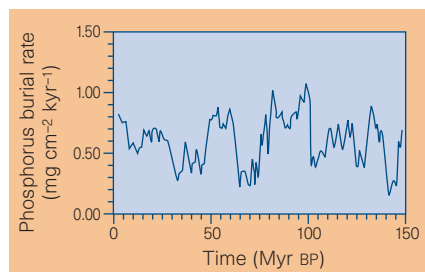


Figure 1 Rate of phosphorus accumulation for all areas and sediment types. The rate is calculated from linear sedimentation rates, dry bulk density (or wet bulk density), porosity and water content⁷.

numerical ages attributed to each measurement using a biostratigraphic distribution chart and the timescale of Harland *et al.*^{7,8}, so it may lack the precision of more recent timescales (especially for the Neogene⁷). However, it covers the whole time span under study, and the trends observed in our data are robust with regard to eventual changes in age. To avoid any subjective bias, we did not modify the timescale⁸.

Figure 1 shows short and long fluctuations in PBR with time that are reminiscent of changes in ocean productivity, circulation pattern and climate; these are intimately linked to the evolutionary history of various processes of the coupled geo-bio-ocean-atmospheric system^{9,10}. Many studies of oceanic geochemical variations have used proxies for phosphorus with the aim of understanding the evolutionary history of ocean biogeochemistry^{9,10}. Recognizing that there are long-term oscillations in the PBR record is essential for understanding exogenic and endogenic forcings.

For an idea of the dominant characteristic timescale of the PBR variability, we used the Thomson multi-taper method (MTM) of spectral analysis, which has low variance, high spectral resolution and accurate statistical significance for the detected spectral peaks^{11,12}. The MTM spectrum of the PBR time series and the corresponding Thomson's statistical reliability test are shown in Fig. 2a, b. These results indicate that there are two consistent peaks centred at 33 ± 3 Myr and 18 ± 1 Myr that are significant at the 95% confidence level. Geophysical and geochemical signals are often dominated by red (low-frequency) background noise, which makes it hard to determine the statistical significance of the spectral components. To find the significance of harmonic components, the MTM adaptive spectra (Fig. 2c) are compared with the robust background-noise spectra¹². Assuming that the ratio of power associated with peaks in the observed continuum spectra to the local power level of background noise follows a chi-squared distribution, we calculated the 95% and 99% confidence levels. The two highly resolved spectral peaks in Fig. 2c are clearly in excess of these significance levels,

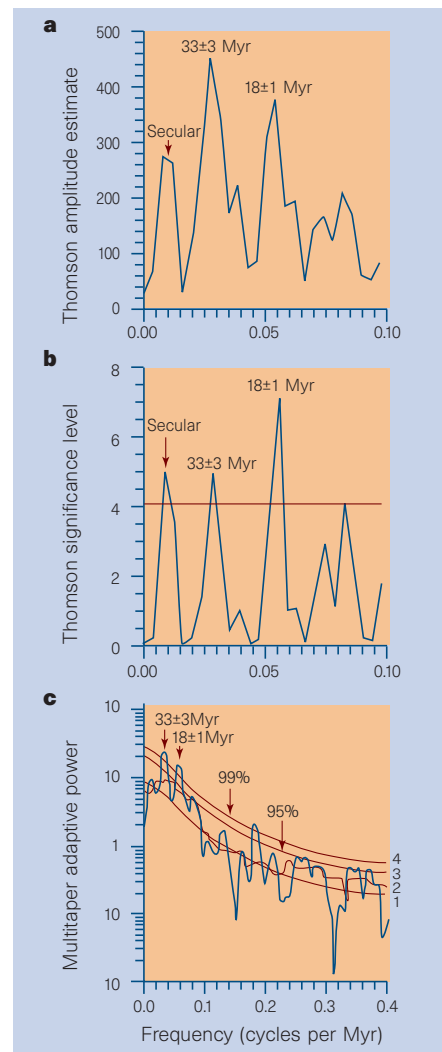


Figure 2 Robust multi-taper spectral analysis. **a**, Thomson multi-taper amplitude spectra. **b**, Thomson's statistical significance *F*-test results (horizontal line indicates 95% confidence level). **c**, High-resolution multi-taper-method adaptive spectra showing: analytical robust background noise spectra curve (1); median smoothed background spectra curve (2); 95% confidence level curve (3); and 99% confidence level curve (4). To compute the adaptive power spectra, we used the following parameters: time-frequency bandwidth $P=2$, and K is the discrete prolate spheroidal wave function (tapers), where $K=2P-1$; for the robust background-noise model, white noise variance¹² $\rho=0.82$ and average power $S_0=1.88$. We used a variable-length moving window for computing median background spectra and found that 0.07 cycle per Myr window is most appropriate.

so we can reject the null hypothesis of a red-noise background continuum.

Earlier claims of a periodicity of 26–33 Myr in terrestrial and extraterrestrial records sparked controversy^{2–4}. Critics of periodicity focused on the incompleteness of the geological records and questioned the statistical reliability of the periodicity detected². Poor geological records and analysis using different spectral techniques might have led to

mismatching of spectral peaks and resolution^{1,4,5}, but we believe that our analysis of a complete and high-resolution record using a powerful spectral technique provides strong evidence of 33-Myr periodicity.

It is not yet clear what drives this periodicity and there is no simple relationship with mass extinction or impact cratering. However, the existence of remarkable spectral power stability and the statistical reliability of our results support the authenticity of this cycle and provide a stimulus for further research into the coupling of bio-geochemical cycles.

R. K. Tiwari, K. N. N. Rao

*Theoretical Geophysics Group,
National Geophysical Research Institute,
Hyderabad 500 007, India
e-mail: postmast@csngri.ren.nic.in*

1. Raup, D. M. & Sepkoski, J. J. *Science* **231**, 833–836 (1986).
2. Stigler, S. M. & Wagner, M. J. *Science* **238**, 940–945 (1987).
3. Raup, D. M. & Sepkoski, J. J. *Science* **241**, 94–96 (1988).
4. Rampino, M. R. & Caldeira, K. *Earth Planet. Sci. Lett.* **114**, 215–227 (1993).
5. Negi, J. G., Tiwari, R. K. & Rao, K. N. N. *Mar. Geol.* **133**, 113–121 (1996).
6. Fischer, A. G. & Arthur, M. A. in *Deep Water Carbonate Environments* (eds Cook, H. E. & Enos, P.) 19–50 (Soc. Econ. Paleont. Mineral., Tulsa, 1977).
7. Follmi, K. B. *Earth Sci. Rev.* **40**, 55–124 (1996).
8. Harland, W. B. et al. *A Geological Time Scale* (Cambridge Univ. Press, 1989).
9. Filippelli, G. M. & Delancy, M. L. *Palaecoenography* **9**, 643–652 (1994).
10. Van Cappellen, P. & Ingall, E. D. *Palaecoenography* **9**, 677–692 (1994).
11. Thomson, D. J. *Proc. IEEE* **70**, 1055–1096 (1982).
12. Mann, M. E. & Lees, J. M. *Clim. Change* **33**, 409–445 (1996).

Life-sustaining planets in interstellar space?

During planet formation, rock and ice embryos of the order of Earth's mass may be formed, some of which may be ejected from the Solar System as they scatter gravitationally from proto-giant planets. These bodies can retain atmospheres rich in molecular hydrogen which, upon cooling, can have basal pressures of 10^2 to 10^4 bars. Pressure-induced far-infrared opacity of H_2 may prevent these bodies from eliminating internal radioactive heat except by developing an extensive adiabatic (with no loss or gain of heat) convective atmosphere. This means that, although the effective temperature of the body is around 30 K, its surface temperature can exceed the melting point of water. Such bodies may therefore have water oceans whose surface pressure and temperature are like those found at the base of Earth's oceans. Such potential homes for life will be difficult to detect.

Planet formation is imperfectly understood, but many models involve the accumulation of solid bodies of up to several Earth masses while the hydrogen-rich solar

nebula is still present¹. These bodies form an envelope of nebula gas bound together by gravity. In the outer Solar System, they may continue to accrete gas, forming the giant planets, while the high ultraviolet output of the early Sun causes those bodies closer to it to lose their gaseous envelope. In one development of these ideas², runaway accretion causes many embryos to form quickly, some of which may merge, while others may be scattered into escape trajectories by proto-Jupiter or proto-Saturn. The formation of planets may be quite inefficient in the sense that more solid material is ejected than retained. There are uncertainties and alternatives^{3–5} but, because solar systems may have been formed in diverse ways, the possibility of bodies of roughly Earth mass in interstellar space should be taken seriously.

The amount of nebular gas accumulated and retained depends strongly on planet mass, nebula temperature, opacity assumptions and accretion timescale. An Earth-mass body eliminating its energy of formation in a million years and with only pressure-induced opacity of hydrogen⁶ develops an atmosphere with $M_{\text{atm}}/M \approx 0.01$, where M is planet mass and M_{atm} is atmospheric mass. More opaque models⁷ yield atmospheric masses with $M_{\text{atm}}/M \approx 0.001$, in agreement with detailed models¹.

The retention of a major part of this atmosphere is difficult at Earth orbit once most of the nebula has cleared, but becomes increasingly likely at greater distances, especially once the atmosphere has cooled (so that the photosphere is no longer large compared with the solid body). The atmospheric escape time can be as short as a million years at one astronomical unit early in the Solar System¹, but longer than the age of the Solar System in the interstellar medium. Sputtering (collision with interstellar molecular or atomic hydrogen at tens to hundreds of kilometres per second) can be important if denser interstellar regions are encountered, but the column density of hydrogen in the case of $M_{\text{atm}}/M \approx 0.001$ to 0.01 is so large that removing such an atmosphere would correspond to much more mass being sputtered per unit area than the total mass per unit area of a comet in the Oort cloud.

At the present epoch (assumed to be around 4.6 Gyr after formation), an interstellar planet would have a luminosity derived from long-lived radionuclides of around $4 \times 10^{20} \chi \text{ erg s}^{-1}$ if it is like Earth⁸, where χ is the planet mass in units of Earth masses. Assuming a thin atmosphere and an Earth-like density, the effective temperature T_e of the planet is given by $T_e \approx 34 \chi^{1/12} \text{ K}$. From hydrostatic equilibrium, the surface pressure $P_s \approx 10^6 \times M_{\text{atm}}/M$ bars. However, optical-depth unity at relevant infrared wavelengths (about 100 μm) is

achieved in such an atmosphere at a pressure of around 1 bar (refs 6,9) and liquefaction at this pressure occurs at a temperature of around 22 K, below the actual atmospheric temperature. A convective gas adiabat must form at all greater depths (at a pressure between 1 bar and P_s), even when the heat flow is very low. An adequate estimate for this adiabat turns out to be $T \propto P^{0.36}$, which does not intercept the condensation curve for hydrogen. It follows that the surface temperature is given by $T_s \approx 425 \chi^{1/12} ((M_{\text{atm}}/M)/0.001)^{0.36} \text{ K}$.

The melting point of water is typically exceeded for basal pressures of the order of one kilobar. The atmosphere will have several cloud layers (methane, ammonia and perhaps water, like Uranus), but this has little influence on the temperature estimate.

It seems, then, that bodies with water oceans are possible in interstellar space. The ideal conditions are plausibly at an Earth mass or slightly less, similar to the expected masses of embryos ejected during the formation of giant planets. Bodies with Earth-like water reservoirs may have an ocean underlain with a rock core. Either way, these bodies are expected to have volcanism in the rocky component and a dynamo-generated magnetic field leading to a well developed (very large) magnetosphere. Despite thermal radiation at microwave frequencies that corresponds to the temperatures deep within their atmospheres (analogous to Uranus⁹), and despite the possibility of non-thermal radio emission, they will be very difficult to detect.

If life can develop and be sustained without sunlight (but with other energy sources, plausibly volcanism or lightning in this instance), these bodies may provide a long-lived, stable environment for life (albeit one where the temperatures slowly decline on a billion-year timescale). The complexity and biomass may be low because the energy source will be small, but it is conceivable that these are the most common sites of life in the Universe. Details of the above results are available from the author.

David J. Stevenson

*Division of Geological and Planetary Sciences,
California Institute of Technology,
Pasadena, California 91125, USA
e-mail: djs@gps.caltech.edu*

1. Hayashi, C., Nakazawa, K. & Nakagawa, Y. in *Protostars and Planets II* (eds Black, D. C. & Matthews, M. S.) 1100–1153 (Univ. Arizona Press, Tucson, 1985).
2. Lissauer, J. J. *Icarus* **69**, 249–265 (1987).
3. Levison, H. F., Lissauer, J. J. & Duncan, M. J. *Astron. J.* **116**, 1998–2014 (1998).
4. Cameron, A. G. W. *Meteoritics* **30**, 133–161 (1995).
5. Boss, A. P. *Science* **276**, 1836–1839 (1997).
6. Birnbaum, G., Borysow, A. & Orton, G. S. *Icarus* **123**, 4–22 (1996).
7. Stevenson, D. J. *Planet. Space Sci.* **30**, 755–764 (1982).
8. Stacey, F. D. *Physics of the Earth* 3rd edn (Brookfield, Brisbane, 1992).
9. Conrath, B. J. et al. in *Uranus* (eds Bergstrahl, J. T., Miner, E. T. & Matthews, M. S.) 204–252 (Univ. Arizona Press, Tucson, 1991).

At the third stroke, the time will be 2300

The Clock of the Long Now

by Stewart Brand

Weidenfeld & Nicholson: 1999. 190 pp.
£12.99

John L. Casti

Daniel Hillis, developer of the Connection Machine parallel computer and founder of The Thinking Machines Corporation, has a new project: a gigantic mechanical clock, perhaps as large as Stonehenge, to be built in the American desert. This fantastic clock is designed to record time for 10,000 years, by ticking (or is it tocking?) once each year, bonging once a century, with the cuckoo coming out on the millennium. The very idea of this clock is to force people's thinking towards long-term awareness and responsibility. As described by Stewart Brand in this provocative volume on long-term responsibilities, "Such a clock ... would embody deep time for people".

The Clock of the Long Now is dedicated to the notion that human civilization has locked itself into a dangerously short attention span. Originating with the ever-accelerating pace of technological change, coupled with the quarterly-report mentality of financial markets and the next-election priorities of democracies, some type of redressing of this pathological imbalance is desperately needed. The story told here is an attempt to correct that short-term mentality by teaching us to take the long view — where 'long-term' is measured in decades and centuries.

In a series of 25 short chapters, none of which is longer than about eight pages, Brand takes us on a whirlwind tour of futuristic technology and philosophical enquiry into our relationship to time. All this is focused on how we might re-orientate our view of time to produce a robust and adaptable civilization. To this end, the author proposes six significant levels of pace and size in the way such a society would work. These levels, from fast to slow, are: fashion/art, commerce, infrastructure, governance, culture and nature.

Brand argues that, in a healthy society, each level is allowed to operate at its own pace, protected from below by slower levels and kept invigorated by those faster levels above. So, for instance, if governance changes suddenly instead of gradually, you get the disasters of the French and Russian Revolutions. But if it changes slowly, you have the more benign American Revolution. By the same token, the job of fashion and art is to be quick, engaging and self-absorbed. In short, the fast layers innovate; the slower layers stabilize. And the whole combines learning and creativity with continuity. This is way things *should* work.

A good deal of Brand's book is about The Clock of the Long Now, which, together with its associated 10,000-Year Library, could help us direct our thinking to these various layers of time. In many ways, I found the 10,000-Year Library even more fascinating than Hillis's Clock, perhaps because I write books and don't design computers. Brand gives an exciting account of what such a library might be good for. It would provide, embody even, the long view of things. Such a library would conserve the information that's needed from time to time for a renaissance. It would help make the world safe for rapid change, since it would give assurance that everything that might need to be remembered is being collected and stored. So if we head down a blind alley or get lost, we can always go back to where we started and try again.

The final chapter of Brand's call-to-the-future centres on James Carse's distinction between a finite and an infinite game — and rightly so, since finite games focus on how they end, while infinite games focus on how they continue. And it's continuation we're interested in when we take the long-term view of things. Continuity is king, not revolution. Infinite games are corrupted by inappropriate finite play, for example when governance (infinite) is disabled when factional combat (finite) becomes the whole point, instead of a basis for constructive debate and alternative modes of power. Similar arguments can be applied to situations when one

culture tries to eradicate another, or when nature is disrupted by commercial competition. The horrors of Kosovo or the logging of the Amazonian rainforest are specific instances of these general dictums.

In one way or another, everyone has a stake in the future. This well-written, interesting, intelligent book is about as good an operating manual as you'll find on how to ensure that that future doesn't slip away through misadventure, miscalculation, or just plain neglect. □

John L. Casti is at the Santa Fe Institute, 1399 Hyde Park Road, Santa Fe, New Mexico 87501, USA.

Wise old wives

Honey, Mud, Maggots and Other Medical Marvels: The Science Behind Folk Remedies and Old Wives' Tales

by Robert and Michèle Root-Bernstein

Macmillan: 1999. 279 pp. £12.99

Rima D. Apple

Medical science often dismisses the knowledge of indigenous peoples, folklore and homespun knowledge as tangential or irrelevant to the advance of medical practice. Yet Robert and Michèle Root-Bernstein's book provides many examples of how such lay wisdom has informed and continues to shape cutting-edge medical practices.



As old as the stones

Was Stonehenge really dragged into place by teams of sweating labourers some 4,600 years ago? Archaeologist Aubrey Burl tackles this classic mystery and comes up with an easier

mode of transport — glaciation — as part of a beautifully illustrated study of legend and history: *Great Stone Circles: Fables, Fictions, Facts* (Yale University Press, \$30, £19.95).

Graphic descriptions of these events track the evolutionary development of medical practice, highlighting areas where modern clinical practice is indebted to 'ancient' medicine, 'foreign' (non-Western) health-care techniques or traditional practices handed down within families over the generations.

At the same time, the stories demonstrate how "Western ways of knowing" made it difficult for many physicians to accept these treatments easily.

Take one of the most famous examples: Edward Jenner's 'discovery' of vaccination. Smallpox killed many and left survivors with ghastly scars and even blind. 'Variolization' (scratching the skin of a healthy person with pus taken from a patient with a mild case of smallpox) was introduced into Western medicine from the East in the early eighteenth century. It usually resulted in only a mild case of the disease and protected the patient from more virulent strains.

Although less dangerous than contracting smallpox naturally, variolization could kill or scar. English dairy farmers recognized a different and less threatening preventative against smallpox: cowpox, which affects humans only slightly. Rural people insisted that people who had had cowpox could safely nurse smallpox patients.

Physicians rejected this idea as foolishness — most physicians, that is. Jenner infected cowpox survivors with weakened strains of smallpox. During the 1796 outbreak of cowpox, Jenner used pus from the arm of a cowpox patient to inoculate a young boy, James Phipps. A few weeks later he inoculated James with smallpox pus. The boy remained immune. Analytically considered, the practices of farmers and milkmaids changed the face of medical practice.

In examining the curative and preventative treatments that have entered clinical practice from non-traditional sources, the authors, perhaps predictably, call for a sense of humility. As the Root-Bernsteins report, people survived before the advent of antibiotics; maggots were and are used successfully to clean wounds and honey solutions were and are used to pack burns. Many cultures around the globe have engaged in mud- and clay-eating; today, we enjoy the benefits of calcium carbonate and kaopectate.

The practice of blood-letting predates written language; contemporary health care continues to recognize the benefits of phlebotomy in conditions such as haemochromatosis, and leeches have once again joined the medical arsenal. Saliva and urine have been used as therapeutic agents over the ages; patients today are rediscovering the antibacterial benefits of these bodily secretions. The historical development of serendipitous and innovative technological interventions, such as cellophane bandages and 'pneumatic leeches', also receive their due in this enter-



Suck it and see: leech therapy from Aldebrando da Firenze's *Medical Treatise*, 1356.

taining and informative book. Written from the perspective of Western medical thought, these stories of unusual therapies are often followed by a contemporary medical explanation of the mechanisms underlying the treatment.

The authors illustrate the evolutionary aspects of medical practice by drawing on connections between therapies observed among other mammals and those practised by humans, such as the licking of wounds and the ingestion of urine. They also investigate the role of culture and economics in the acceptance or rejection of medical therapeutics, recognizing that "the best treatment is not simply the most efficacious one, but the one that is actually put to use".

Their conclusion delineates the very real economic barriers to further incorporation of folk medicine into our current health-care system. In most countries today, medical therapies must be approved by a regulatory body before they can be marketed or distributed. For example, in the United States it costs millions of dollars to fund the clinical trials demanded by the approval process of the Food and Drug Administration, not to mention millions for product development, manufacture, advertising and the like.

Such an enterprise would be commercially unfeasible for an inexpensive, readily available product such as honey that cannot be patented. Moreover, despite their generally positive portrayal of non-traditional therapies, towards the end of the book, the Root-Bernsteins issue a warning against cavalierly embracing all such modalities, and provide numerous examples of health-care fraud based on anecdotal evidence from folk medicine. But this short chapter offers little concrete advice on how non-medical people can make meaningful choices.

Still, this entertaining book is replete with fascinating tales examining the untraditional, at times irrational, practices that have been adopted by contemporary medical practice. □

Rima D. Apple is at the School of Human Ecology, University of Wisconsin-Madison, Madison, Wisconsin 53706, USA.

How far the stars?

Measuring the Universe: The Historical Quest to Quantify Space

by Kitty Ferguson

Headline: 1999. 306 pp. £14.99

Virginia Trimble

From Earth and Moon to Sun and stars and on to galaxies and the Universe, our view of the world has grown more or less monotonically throughout recorded history (at least Western history). It is this expanding horizon that Kitty Ferguson has set out to map. Her chosen method is to focus on specific measurements that yielded numbers we now regard as accurate, from Eratosthenes and the diameter of the Earth to current attempts to pin down the cosmic expansion rate (Hubble parameter), the density of the Universe and the infamous cosmological constant. Ferguson tries to give the reader a feeling for what it must have been like to attempt these measurements for the first time.

I came to the book thinking the task worth doing and worth doing well, and so was prepared either to admire elegant explanations of difficult concepts or to carp unmercifully at errors, omissions and misconceptions. Sadly, the negative points greatly outweigh the positive ones. My list of 'bravos' has half-a-dozen items, including an appealing preface on triangulating windmills, a good epilogue, a thought-provoking quote from Fred Hoyle on the impossibility of knowing in advance which problems are solvable by current methods, and an excellent approach to acronyms: she uses very few.

On the other side of the balance sheet: my list of 'truly wrong' has 40 items; 'misleading' = 48; 'something vital left out' = 45; 'poorly explained, perhaps because not understood' = 52; and other miscellaneous oddities, including sudden switches of topic in the middle of a story = 29.

How much do these things matter? Some hardly at all, unless you are personally concerned or need to tell others about the topic. Consider the following (and only) description of the discovery of quasars: "Sandage was involved in some of the first investigations of these objects in the early 1960s, along with Thomas Matthews and Maarten Schmidt." This is, I believe, unique in the literature in the extent to which it understates the contribution of Schmidt, the actual discoverer, although, as a corrective to common overstatements, it also fails in not crediting the radio astronomers, without whose accurate coordinates nothing would have happened.

Somewhat more serious are omissions that leave a very incomplete picture. For instance, neutrinos of non-zero rest mass are the only potential form of dark matter pre-

sented in a chapter called "The quest for omega", and they are credited to very recent laboratory results. Where are the cold dark matter candidates like Weakly Interacting Massive Particles and axions; where the baryonic candidates (MACHOs and brown dwarfs), strings, and all the rest? Where, for that matter, is the 20-year history of neutrinos as dark matter? Nowhere. And, indeed, omega is nowhere defined.

The intended audience consists of the customary intelligent laymen and women (the book is generally careful about such matters), with zero knowledge but infinite capacity. These probably mythical persons will not notice and may not care about numerical and historical errors and omissions. But incomprehensible explanations are surely fatal. Said readers will be left feeling either unintelligent or mistreated. Thus, I suspect they will cast the book aside about the time (*circa* 300 BC or the middle of Chapter 1) when Aristarchus of Samos is trying to learn the distances to the Moon and Sun.

Aristarchus' attempt to get the solar distance from the unequal lengths of the four main phases of the Moon (from new to first quarter, first quarter to full, and so on) is extraordinarily clever. Why did he fail? The Moon's orbit is an ellipse, not a circle, but Ferguson never tells you this. The figures do not seem to help difficult geometric concepts; author and illustrator conspire together

to keep the reader from discovering that, if you can measure two angles of a triangle and the included side, you have the rest.

Modern territory brings us to ways of measuring distances to galaxies, clusters and quasars very far from our own Milky Way. Those involving sources of known absolute brightness come across fairly well, but the more complex ones could not be understood from what is given even by infinite intelligence. You always need to measure two things (and often do some modelling as well). For distances to gravitationally lensed quasars, the two things are the time delay between flares seen in separate images and their angular separation on the sky. Only the time delay is mentioned. For the hot gas in rich clusters of galaxies, the two are the degree to which gas scatters photons of the cosmic microwave background (for which the author gets the sign wrong) and its X-ray emission. Only the scattering is mentioned.

Many popular books of popular science, both in and out of my field of astronomy, have passed the test: "Would you be reading this if you hadn't been asked to review it?", including ones by Fred Hoyle, Carl Sagan, Isaac Asimov, Peter Medawar, Malcolm Longair, Martin Rees, Walter Sullivan and Stephen Jay Gould. *Measuring the Universe* is not one of these books. □

Virginia Trimble is in the Department of Physics, University of California, Irvine, California 92697-

4575, and in the Astronomy Department, University of Maryland, College Park, Maryland 20742, USA.

Further astronomy

Astronomy in Prehistoric Britain and Ireland

by Clive Ruggles

Yale University Press, \$65, £45

Sustaining hopes for change

The Brundtland Commission's Report – 10 Years

edited by Guri Bang Sjøfting, George Benneh, Kjetil Hindar, Lars Walløe and Anders Wijkman

Scandinavian University Press/Research Council of Norway: 1998. 234 pp. Nkr298

Peter Doran

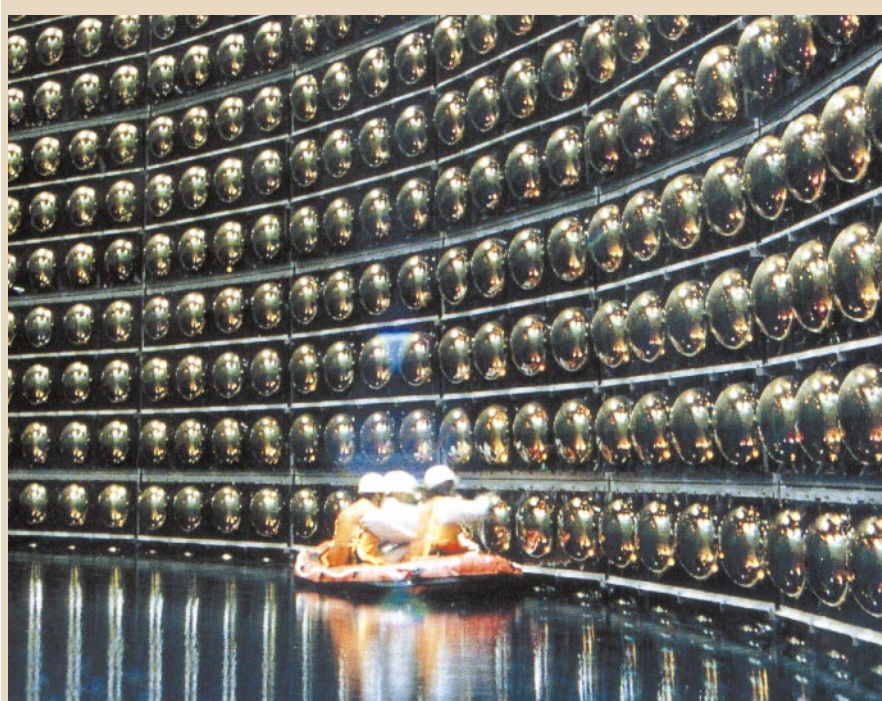
More than ten years after the landmark publication of *Our Common Future* by the World Commission on Environment and Development (WCED), the questions that face the growing community of policy-makers and researchers in sustainable development have changed radically.

Research agendas are no longer driven simply by a need to understand the various categories and causes of unsustainability in our social, economic and environmental decision-making. They are increasingly concerned with knowledge itself and reflect the important challenge of making better use of existing information and framing that knowledge in a way that makes it useful and accessible at all levels of decision-making.

At the local level, for example, research that leads to understanding and action must often be pursued with the participation of local people. Such approaches to knowledge-gathering most frequently occur where empowerment, motivation, understanding and democracy have become interlinked. The Nordic countries' enthusiastic response to the sustainable-development movement seems to support the view that a democratic culture and an informed citizenry provide the most fertile soil for new values and behaviour patterns to take hold.

These changes to the research agenda for sustainable development are reflected in the broad recommendations in *The Brundtland Commission's Report*. This is a collection of papers produced for an international research conference in Oslo to mark the tenth anniversary of the publication of the WCED report. The conference was designed to take stock of the "evolution of sustainable development up to the present", and the Norwegian Research Council is already planning a research programme around some of the points that came out of it.

Blessed with an optimism that can



Underground enlightenment

Technicians travel by raft under the giant dome of the Kamiokande neutrino detector, a kilometre under a mountain near Tokyo, to clean the faces of photomultipliers. These detect the presence of neutrinos by the light created

when they collide with atoms. This is just one of 220 mind-expanding illustrations — with clear, informative text about the advances physics has made this century — in *Physics in the Twentieth Century* by Curt Supplee (Abrams, \$49.50, £32).

accompany a world view formed by the Scandinavian experience, the editors attribute the popularization of the term “sustainable development” to the work of the WCED, and suggest that it is “leading to a historic shift in all of our societies”. Scandinavian countries have certainly taken a lead in championing special issues raised in the aftermath of the 1992 United Nations Conference on Education and Development. Finland, for example, has undertaken an excellent research programme to investigate the implications of sustainable development on a social dimension. Norway has taken a special interest in the issues of production and consumption.

In her contribution to the volume, Gro Harlem Brundtland, who chaired the commission named after her, explains how a transformation in the political climate in Western countries has helped to extend this change in the fortunes of the sustainable-development agenda beyond Scandinavia. The WCED reported during a period of political conservatism and economic liberalism in Western countries, and few governments supported the calls for stronger governance in their economies. The recent ascendancy of social democratic-led coalitions, in some cases including Green Party representatives, marks a dramatic shift in Europe’s political and democratic landscape, and transforms the prospects for policies of sustainable development.

These prospects are also investigated in the context of the developing countries. Madhav Gadgil observes that, while the globalization of the flow of goods and services has not reached poorer countries, better access to information can make it easier for ordinary people to be involved in planning and managing change, contributing to a shift from control-and-command structures to a culture of inform-and-share.

Gadgil cites the example of the Kerala Sastra Sahitya Parishat (KSSP), a popular-science movement in India which focuses on health and environment issues and invites ordinary people to question accepted patterns of development.

The importance of institutionalizing new approaches to knowledge and learning by society is a recurring theme of this volume, which is a galaxy of potential thesis topics and book projects.

In a section on production and consumption, John R. Ehrenfeld highlights the importance of indicators as a means of maximizing the chance that societies will be able to make the necessary changes. Such indicators, together with institutionalized learning, he argues, are essential if sustainable patterns of production and consumption are to be achieved. □

Peter Doran is in the Department of Humanities, Magee University College, Northland Road, Derry BT49 7JL, Northern Ireland, UK.

Science in culture

Sexy stamens and provocative pistils

The taxonomic science of the flower bed
Martin Kemp

For centuries the study of flowers and the cultivation of gardens were deemed to be safe pursuits for young ladies. The behaviour of animals, by contrast, was all too likely to provoke difficult questions about sexual activity. Carl Linnaeus’s sexual system for the classification of plants, based on stamens and pistils, changed all that.

Introduced to a worldwide readership in his *Philosophia Botanica* of 1751, Linnaeus’s principles attracted fervent adherents and keen opposition. Among the devotees was Erasmus Darwin, Charles’s grandfather, who was an enthusiast of the French Revolution and adopted a radically libertarian stance on social matters. Erasmus’s scientific poem, “The loves of plants”, published in 1789 as part II of *The Botanic Garden*, blends sober scientific analysis with poetic rapture, the latter typified by his evocation of the polygamy practised by *Gloriosa superba*:

*Proud Gloriosa led by three chosen swains,
The blushing captives of her virgin chains...
When time’s rude hand a bark of wrinkles spread
Round her limbs, and silver’d o’er her head,
Three other youths her riper years engage,
The flatter’d victims of her wily age.*

If this should sound like a perversion of Linnaeus’s method, we may recall that the great Swedish botanist had written that “The flower’s leaves ... serve as bridal beds which the creator has so gloriously arranged ... and perfumed with so many soft scents that the bridegroom with his bride might there celebrate their nuptials with so much greater solemnity. When now the bed is so prepared, it is time for the bridegroom to embrace his beloved bride and offer her his gifts.”

Unsurprisingly, religious and conservative organizations began to express alarm. Most notably, the *Encyclopaedia Britannica*, published from the Calvinist redoubts of Edinburgh in 1768, railed against the “disgusting strokes of obscenity” with which Linnaeus had disfigured the picture of nature’s innocent beauties.

The illustrated book that best captured the tone of Darwinian rapture was Robert Thornton’s majestic but failed enterprise, the *New Illustration of the Sexual System of Linnaeus*, first advertised to subscribers in 1797 and appearing in parts from 1799. The illustrations ranged from tabular and diagrammatic representations of Linnaeus’s system, to romantic pictures of particular plants in evocative landscapes and highly charged allegories of nature. For example, *Strelitzia reginae* or ‘queen plant’, named after Charlotte of Mecklenburg-Strelitz, George III’s queen, is soberly anatomized in one plate, depicted in an



Philip Reinagle, “Cupid Inspiring the Plants with Love”, from Robert Thornton’s *Temple of Flora*, 1804, plate III.

exotic setting in another, and features as the target of Cupid’s arrow in the allegorical image of “Cupid inspiring the plants with love” from the pictorial part of the *New Illustration* — re-titled in 1804 as *The Temple of Flora*.

However, lest we should think that Queen Charlotte was being encouraged to distribute her favours with Darwinian profligacy, Thornton is at pains to eulogize the royal patron of his enterprise as a “bright example of conjugal fidelity and maternal tenderness”. Unshakeably pious, staunchly monarchist and very English, Thornton would have nothing to do with Erasmus Darwin’s dangerously French attitudes.

Thornton argued that the mathematical, logical character of taxonomy was a “noble exercise” which was eminently suitable for the training of the minds of the young, who are all too easily seduced by pastimes that “inflamm[e] the passions”. By stressing the dispassionate character of classification, he was consciously confronting the accusation that the sexual basis of Linnaeus’s method was an obscene perversion of the innocence of plants and besmirched botany as a study unfit for young ladies. Thornton was determined that the Linnaean binomial system should serve as the taxonomic science of the flower bed and not as a justification for abandoning the proper regulation of the human nuptial chamber. □

Martin Kemp is in the Department of the History of Art, University of Oxford, 35 Beaumont Street, Oxford OX1 2PG, UK.

Sister-chromatid separation at anaphase onset is promoted by cleavage of the cohesin subunit Scc1

Frank Uhlmann, Friedrich Lottspeich* & Kim Nasmyth

Research Institute of Molecular Pathology, A-1030 Vienna, Austria

* Max Planck Institute of Biochemistry, 82152 Martinsried, Germany

Cohesion between sister chromatids is established during DNA replication and depends on a multiprotein complex called cohesin. Attachment of sister kinetochores to the mitotic spindle during mitosis generates forces that would immediately split sister chromatids were it not opposed by cohesion. Cohesion is essential for the alignment of chromosomes in metaphase but must be abolished for sister separation to start during anaphase. In the budding yeast *Saccharomyces cerevisiae*, loss of sister-chromatid cohesion depends on a separating protein (separin) called Esp1 and is accompanied by dissociation from the chromosomes of the cohesin subunit Scc1. Here we show that Esp1 causes the dissociation of Scc1 from chromosomes by stimulating its cleavage by proteolysis. A mutant Scc1 is described that is resistant to Esp1-dependent cleavage and which blocks both sister-chromatid separation and the dissociation of Scc1 from chromosomes. The evolutionary conservation of separins indicates that the proteolytic cleavage of cohesin proteins might be a general mechanism for triggering anaphase.

The separation of sister chromatids at the metaphase-to-anaphase transition is one of the most dramatic events of the eukaryotic cell cycle. As cells enter mitosis, chromosome condensation during prometaphase resolves the bulk of each chromatid's chromatin from that of its sister^{1,2}. Chromatids nevertheless remain paired along their entire length during the attachment of chromosomes to the mitotic spindle. Cohesion between sisters resists the pulling forces exerted by microtubules attached to sister kinetochores³ and thereby ensures that sister chromatids attach to microtubules emanating from opposite spindle poles^{4,5}. It has long been suspected that destruction of sister-chromatid cohesion, rather than a major change in traction exerted by the spindle, is responsible for the sudden separation of sister chromatids at the metaphase-to-anaphase transition^{3,4}. It is not known what triggers this event in the eukaryotic cell cycle.

There are important clues as to the molecular nature of the cohesive structures that holds sisters together and the mechanism by which it is suddenly broken at the onset of anaphase⁶. In *S. cerevisiae*, cohesion between sister chromatids depends on a multisubunit complex, called cohesin, which contains at least four subunits: Scc1, Scc3, Smc1 and Smc3 (refs 7–9). Cohesion is established during DNA replication with the help of Scc2 and Eco1 (refs 9–11). A similar cohesin complex has been implicated in sister-chromatid cohesion in *Xenopus* extracts².

In yeast, there is a sudden change in the state of cohesin at the metaphase-to-anaphase transition: two cohesin subunits, Scc1 and Scc3, suddenly disappear from chromosomes at the point when sister chromatids separate^{7,9}. The dissociation of Scc1 from chromosomes and the separation of sister chromatids both depend on a 'separin' protein called Esp1 (ref. 12). The existence of Esp1 homologues in many eukaryotes, including humans, suggests that separins have a fundamental and conserved role in chromosome segregation^{13–15}.

For much of the cell cycle, Esp1 is tightly bound by the anaphase inhibitor Pds1 (ref. 12), whose destruction at the metaphase-to-anaphase transition is triggered by ubiquitination due to the anaphase-promoting complex (APC)¹⁶. The APC requires an

activator protein, Cdc20, to mediate Pds1 destruction¹⁷. In *S. cerevisiae*, the only role of the APC in promoting sister separation is to destroy Pds1 (refs 12, 18). We now investigate the mechanism by which Esp1 dissolves sister-chromatid cohesion once it has been liberated from Pds1.

Esp1 controls Scc1 chromosome association

The displacement of Scc1 from chromosomes at the metaphase-to-anaphase transition might be a direct effect of Esp1 activity. Alternatively, it might just be a consequence of sister-chromatid separation initiated by Esp1. We therefore examined the mechanism that prevents the association of Scc1 with chromosomes during early G1 phase. Scc1 is destroyed during anaphase and is normally not resynthesized until late G1 in the next cell cycle⁷. However, even when Scc1 is synthesized in early-G1 cells from the galactose-inducible *GAL1-10* promoter, it fails to bind stably to chromosomes¹⁰. We repeated this experiment using cells arrested in a G1-like state with the mating pheromone α -factor (Fig. 1a). Scc1, induced during pheromone arrest, accumulated in the nuclei of most cells, but bound to chromosomes only weakly, if at all. Furthermore, the protein rapidly disappeared from cells after expression was shut off by addition of glucose (Fig. 1b). When the experiment was repeated with *esp1-1* mutant cells, Scc1 bound to chromosomes with high efficiency and remained associated even after synthesis was terminated (Fig. 1b). This result indicates that Esp1 prevents the stable association of Scc1 with chromosomes during G1 phase, in addition to causing dissociation of Scc1 from chromosomes when sister chromatids separate. Esp1 may therefore have a direct role in removing Scc1 from chromosomes.

An *in vitro* Scc1-dissociation assay

To investigate the mechanism by which Esp1 causes Scc1 to dissociate from chromosomes, we assayed this process *in vitro* (Fig. 2). A crude preparation of yeast chromatin¹⁹, isolated from cells arrested in a metaphase-like state by nocodazole, was incubated with soluble extracts from *esp1-1* mutant cells that either had or had not been induced to overexpress wild-type Esp1 from the *GAL1-10*

promoter. After incubation with both types of extract, the chromatin fraction was again separated from the supernatant by centrifugation and the levels of haemagglutinin(HA)-epitope-tagged Scc1 in chromatin and supernatant fractions were analysed by SDS-PAGE and subsequent immunoblotting. About 70% of the total Scc1 in nocodazole-blocked cells is tightly associated with chromatin⁹ and is therefore present in the starting chromatin fraction, which served as substrate. Most Scc1 remained in the chromatin fraction following incubation with extract prepared from *esp1-1* mutant cells, but almost all disappeared from the chromatin fraction after incubation with extract containing overexpressed Esp1 (Fig. 2). Surprisingly, Scc1 induced to dissociate from chromatin by Esp1 appeared in the supernatant fraction as a cleaved product (Fig. 2). Both dissociation of Scc1 from chromatin and its cleavage were inhibited by Pds1 translated in reticulocyte lysate, but not by a control lysate (Fig. 2). We also detected a small amount of Scc1-cleavage activity in extracts prepared from cells not overexpressing Esp1 if they were arrested in G1, and in extracts from cycling cells lacking Pds1 (data not shown). Esp1 also caused ~50% of the Scc3 cohesin subunit⁹ to dissociate from chromatin without cleavage. The association of Smc proteins and histone H2B1 with chromatin was unaffected by Esp1-containing extracts (data not shown).

We next investigated the requirements for Scc1 dissociation *in vitro*. It has been suggested that a transient calcium wave in mitotic cells might trigger sister-chromatid separation²⁰, but Scc1 cleavage was unaltered by addition of the calcium chelator EGTA or by an excess of free calcium. The reaction was also not inhibited by inhibitors of kinases or phosphatases (data not shown). This suggests that the dissociation of Scc1 from metaphase chromatin *in vitro* is neither induced by calcium nor by *de novo* phosphorylation/dephosphorylation. ATP depletion of extracts or addition of the proteasome inhibitor LLNL did not prevent Scc1 cleavage (data not shown), indicating that Scc1 cleavage is probably not due to APC-mediated ubiquitination. We attempted to characterize the proteolytic activity by using protease inhibitors from the four known classes, and found that Scc1 cleavage was only inhibited by high concentrations (10 mM) of *N*-ethylmaleimide.

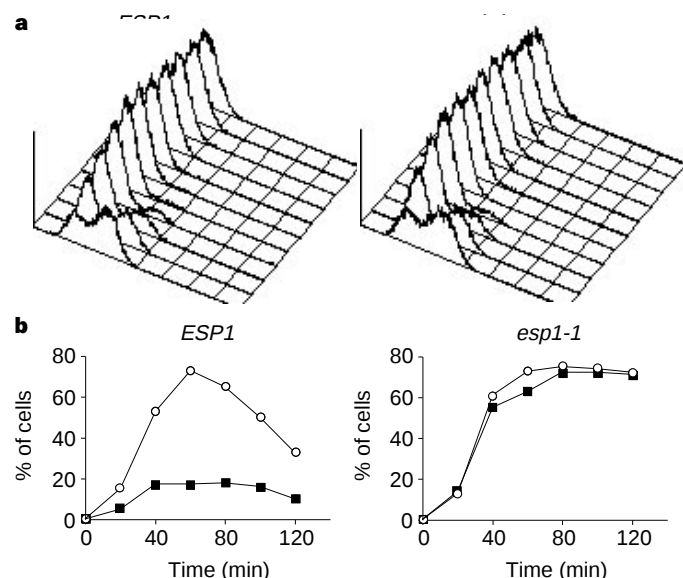


Figure 1 Esp1-dependent chromosome association of Scc1 in G1. **a**, Strains K7466 (*MATa ESP1 SCC1 GAL-SCC1myc18*) and K7468 (*MATa esp1-1 SCC1 GAL-SCC1myc18*) were arrested in G1 with α -factor for 120 min (time point zero). FACS analysis showed that cells stayed arrested during the experiment. **b**, Scc1-Myc18 was induced for 60 min, then cells were transferred to medium containing glucose to repress Scc1-Myc18. Expression of Scc1-Myc18 was seen by whole-cell *in situ* staining (circles), and chromosome binding of Scc1-Myc18 was observed by using chromosome spreads (squares).

Scc1 cleaved at anaphase onset *in vivo*

To test whether Esp1-induced cleavage of Scc1 also occurs *in vivo* at the onset of anaphase, we used a yeast strain in which expression of the APC-activator Cdc20 is under the control of the galactose-inducible *GAL1-10* promoter²¹. Cells from this strain were arrested in metaphase by incubation in galactose-free medium and then induced to undergo synchronous anaphase by addition of galactose. Sister-chromatid separation, visualized by the binding to *tet* operators close to CenV of *tet* repressor tagged with green fluorescent protein (GFP) (ref. 7), occurred in most cells within 15 min of Cdc20 induction, and Scc1 dissociated from chromosomes at a similar rate (Fig. 3a, b). We detected a small amount of Scc1 cleavage product of the same size as that seen *in vitro* in cycling cells, but not in cells arrested in metaphase. The cleavage product appeared 15 min after release into anaphase, simultaneously with sister-chromatid separation and the dissociation of Scc1 from chromosomes (Fig. 3c). Full-length Scc1 remaining in cells at this time might originate from the soluble pool of Scc1, whose cleavage is unnecessary for sister separation. Soluble Scc1 in the supernatant fraction of chromatin preparations makes a poor substrate in our *in vitro* cleavage assay (data not shown).

To establish whether cleavage of Scc1 during anaphase depends on Esp1, we compared wild-type and *esp1-1* mutant cells after their release from *GAL-CDC20* arrest at 35 °C. The extent of sister-chromatid separation, Scc1 dissociation from chromosomes (data not shown), and Scc1 cleavage (Fig. 3d) was greatly reduced in the *esp1-1* mutant. We conclude that Esp1 promotes cleavage of Scc1 and its dissociation from chromosomes both *in vivo* and *in vitro*.

A cleavage-resistant Scc1

To determine whether Esp1-mediated cleavage of Scc1 is a cause or a consequence of its dissociation from chromosomes, we first identified the Scc1-cleavage site, with a view to producing a cleavage-resistant mutant. The C-terminal Scc1-cleavage product in anaphase cells (Fig. 3) was immunoprecipitated from cell extracts. Amino-terminal sequence analysis of the fragment showed that cleavage had occurred between a pair of arginine

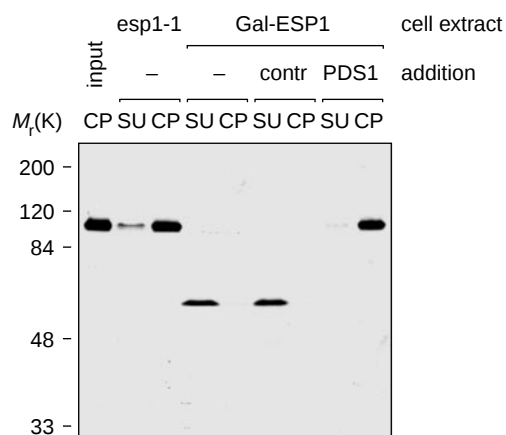


Figure 2 *In vitro* assay for Scc1 dissociation from chromatin. Chromatin was prepared as described¹⁹ from strain K7563 (*MATa SCC1-HA6*) arrested in metaphase by nocodazole treatment. Proteins in the chromatin preparation were resolved by SDS-PAGE; Scc1-HA6 was detected by western blotting (input). This chromatin preparation was resuspended in the indicated extracts, with or without addition of 50% (v/v) *in vitro* translation reactions, as indicated. After incubation, aliquots of the supernatant fraction (SU) and the chromatin fraction (CP) of each reaction were analysed.

residues at positions 268 and 269. The first of these arginine residues was then mutated to aspartic acid (R268D), tagged at the C terminus with HA epitopes, and expressed from the *GAL1-10* promoter: expression of the mutant protein had little effect on cell proliferation. To test whether the R268D mutation had abolished cleavage, we used chromatin from cells expressing it as a substrate in the Esp1 assay. We found that there was no cleavage at site 268, but that the mutant protein was still cleaved in an Esp1-dependent manner (Fig. 4a). The C-terminal cleavage product was now about 10K larger. To identify the second cleavage site, we looked for sequences in Scc1 that are similar to those around the C-terminal cleavage site and found a 5-out-of-7 amino-acid match at position 180 (Fig. 4b). We mutated the arginine before this putative cleavage site to aspartate (R180D).

We next compared the effect of expressing wild-type Scc1, R180D and R268D single-mutant proteins, and the R180D/R268D double-mutant protein from the *GAL1-10* promoter. Neither the single-mutant proteins nor the wild-type protein greatly affected cell proliferation, but expression of the double-mutant protein was lethal (data not shown). We used chromatin from cells transiently expressing the double-mutant protein in our Esp1 assay and found that the R180D/R268D double-mutant protein (Scc1RR-DD) was no longer cleaved. Furthermore, it failed to dissociate from chromosomes (Fig. 4a). The small amount of 'leakage' of Scc1 from chromatin into the supernatant was Esp1-independent (data not shown).

To obtain the N- and C-terminal cleavage products simultaneously, we used as a substrate chromatin from a strain expressing Scc1 tagged N-terminally with a Myc epitope and C-terminally with HA epitope (Fig. 4b). Esp1-mediated cleavage produced a single HA-tagged cleavage fragment but two Myc-tagged fragments, the smaller of which was more abundant (Fig. 4c). These results indicate that all molecules were cleaved at the C-terminal site but not all of them were cleaved at the N-terminal site. A similar pattern of N-terminal cleavage fragments was obtained during anaphase *in vivo* (data not shown).

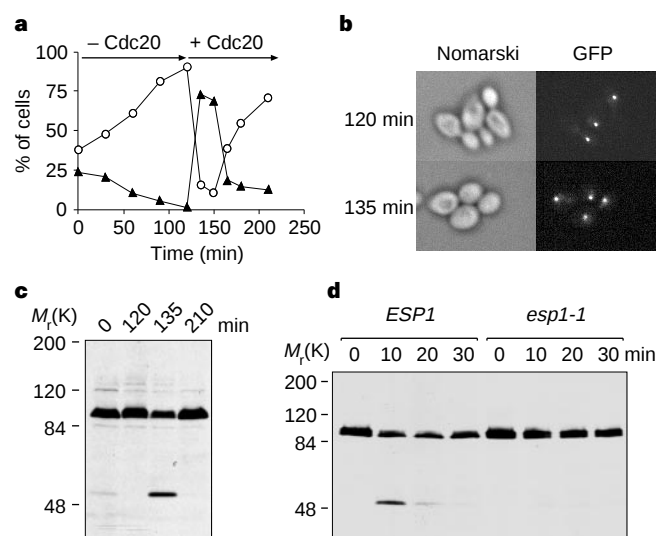


Figure 3 Scc1p cleavage at anaphase onset *in vivo*. **a**, Metaphase arrest and release using Cdc20 depletion of strain K7677 (*MAT α cdc20 Δ GAL-CDC20 Scc1-HA3 TetR-GFP TetOs*). Scc1-HA3 bound to chromosomes (circles), and the fraction of cells containing two separated GFP dots (triangles). **b**, Examples of cells in the arrest at 120 min, and 15 min after release. **c**, Western blot analysis of Scc1-HA3 in whole-cell extracts prepared from cells at the indicated time points. **d**, As **c**, except that strain K7677 and K8054 (*MAT α esp1-1 cdc20 Δ GAL-CDC20 Scc1-HA3 TetR-GFP TetOs*) were arrested and released at 35°C.

Non-cleavable Scc1 and sister separation

To investigate why cells expressing the non-cleavable Scc1RR-DD double mutant cannot proliferate, we used centrifugal elutriation to isolate G1 cells from a culture growing without expression of Scc1RR-DD, which were then incubated in the presence and absence of galactose to induce Scc1RR-DD from the *GAL1-10* promoter (Fig. 5). To minimize the duration of mutant protein expression, cells grown in the presence of galactose were transferred to glucose-containing medium after 135 min, when most cells had replicated their DNA (Fig. 5a). In the absence of galactose, sister separation and the dissociation from chromosomes of endogenous Myc-tagged Scc1 occurred simultaneously about 60 min after DNA replication (Fig. 5b–d). Transient expression of Scc1RR-DD, tagged with HA epitope, almost completely prevented sister-chromatid separation (Fig. 5b) but did not affect binding and dissociation of endogenous wild-type Scc1 (Fig. 5c). The mutant protein remained tightly associated with chromosomes long after the endogenous wild-type protein had disappeared (Fig. 5c, d). Scc1RR-DD bound to chromosomes immediately following induction in G1, when wild-type Scc1 is prevented from binding by Esp1 (compare to Fig. 1 and

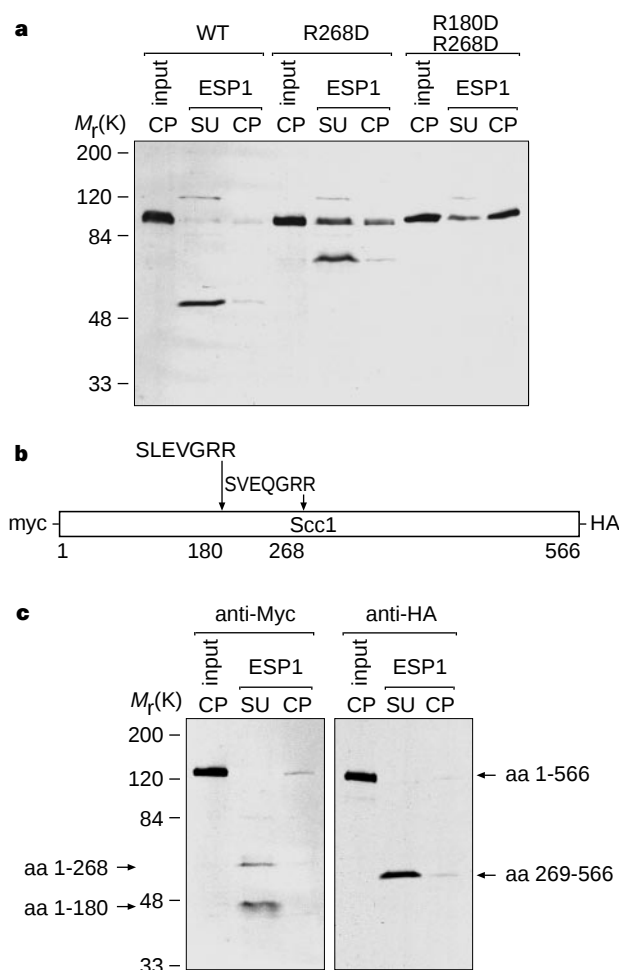


Figure 4 Characterization of the Scc1 cleavage sites. **a**, Strains K8097 (*MAT α Scc1-myc18 GAL-Scc1-HA3 TetR-GFP TetOs*), K8099 (*MAT α Scc1-myc18 GAL-Scc1(R268D)-HA3 TetR-GFP TetOs*) and K8101 (*MAT α Scc1-myc18 GAL-Scc1(R180D, R268D)-HA3 TetR-GFP TetOs*) were grown in medium containing 2% raffinose. Expression of the respective Scc1 variant was induced for 4 h, and cells were arrested by nocodazole treatment. Chromatin was prepared and used in the Esp1 assay. The 120K bands are HA-cross-reacting proteins. WT, wild type. **b**, The cleavage sites in Scc1. **c**, Chromatin was prepared from strain K7768 (*MAT α myc9-Scc1-HA6*). Scc1 (ref. 8) and the derived cleavage fragments (aa, amino-acid residues) migrate abnormally slowly during SDS-PAGE, as shown by western blotting with the indicated antibodies.

ref. 10). The failure of Scc1RR-DD to dissociate from chromosomes was not an artefact due to transient overexpression of the protein from the *GAL1-10* promoter, because wild-type Scc1 expressed similarly dissociates from chromosomes with normal kinetics¹⁰.

Expression of the Scc1RR-DD mutant protein caused a transient delay of cytokinesis for 20 min (Fig. 5b), after which cells divided

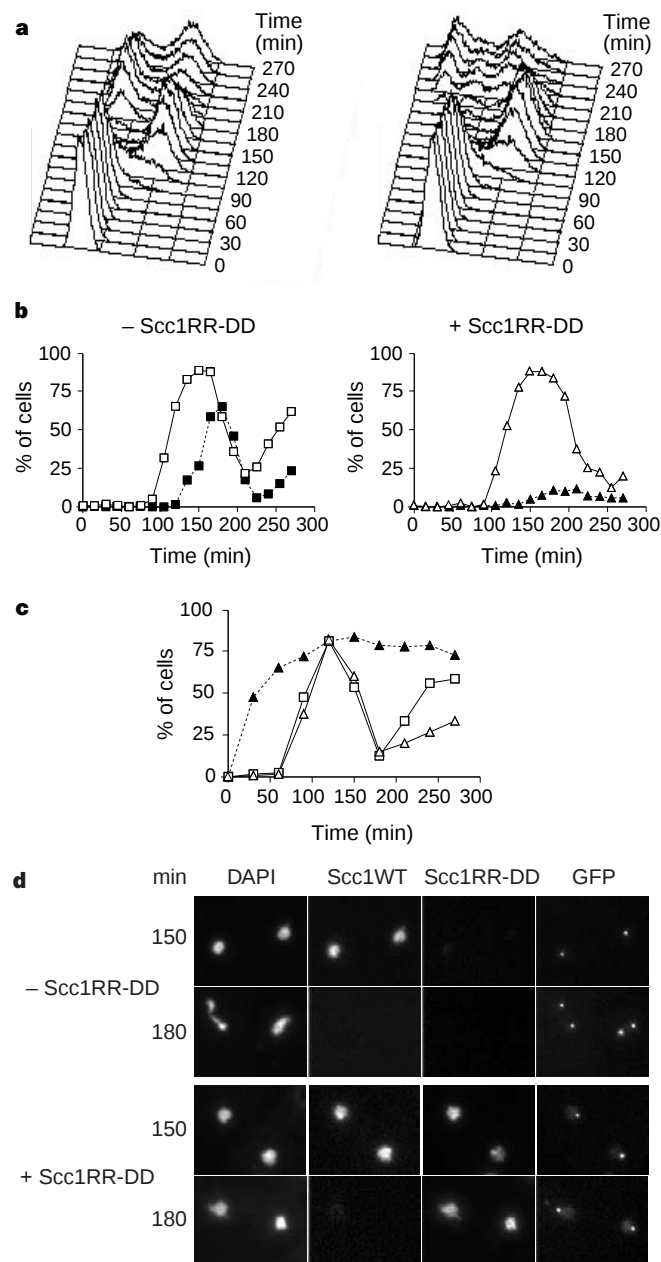


Figure 5 Expression of non-cleavable Scc1 prevents sister chromatid separation. **a**, DNA content as unbudded G1 cells of strain K8101 were released into the cell cycle with or without induction of the Scc1RR-DD mutant. **b**, Budding index (without Scc1RR-DD, open squares; with Scc1RR-DD, open triangles) and percentage of cells with separated sister chromatids (filled symbols). **c**, Scc1 chromosome association. Endogenous wild-type Scc1-Myc18 in cells without (squares) and with Scc1RR-DD expression (open triangles). Scc1RR-DD was tagged with HA epitopes (filled triangles). **d**, Examples of chromosome spreads at 150 min in metaphase and at 180 min when most cells of the control culture had undergone anaphase. DNA was stained with DAPI, Scc1-Myc18 was detected with rabbit anti-Myc antiserum and Cy5-conjugated secondary antibody, Scc1RR-DD-HA3 was detected with mouse monoclonal antibody 16B12 and Cy3-conjugated secondary antibody. Sister chromatids of chromosome V were visible by the GFP dots.

without having separated sister chromatids. Progeny with abnormal DNA contents were produced (Fig. 5a), resembling *esp1-1* mutant cells incubated at the restrictive temperature²². The dissociation from chromosomes of wild-type protein on cue shows that Esp1 activity was not impaired in these cells. The failure of sister chromatids to separate even when Esp1 was active also prevented elongation of the mitotic spindle (data not shown), consistent with the idea that loss of cohesion triggers anaphase.

To determine the phenotype of single- and double-mutant Scc1 proteins when expressed from the natural *SCC1* promoter, we transferred the mutations to *SCC1* carried on a centromeric vector. Plasmids carrying the wild-type or either of the single-mutant genes transformed wild-type and *scc1-73* strains and complemented the temperature-sensitive phenotype of the *scc1-73* mutation. No transformants expressing the R180D/R268D double mutant were obtained (data not shown). Together, our results indicate that cleavage of Scc1 at one of two sites is necessary both for sister-chromatid separation and for the dissociation of Scc1 from chromosomes.

Non-cleavable Scc1RR-DD is functional

To test whether Scc1RR-DD is fully functional apart from its non-cleavability, we first checked whether the double-mutant protein could establish cohesion by itself between sister chromatids in the absence of endogenous Scc1 function, and then whether cohesion established by Scc1RR-DD was dependent on other cohesin subunits. We used centrifugal elutriation to isolate G1 cells of *scc1-73* and *smc3-42* mutant strains⁷ that could express Scc1RR-DD from the *GAL1-10* promoter, then incubated them in the presence or absence of Scc1RR-DD at 35 °C, a restrictive temperature for both mutations. In the absence of galactose, sister chromatids separated prematurely in both mutants and failed to segregate to opposite poles of the cell⁷ (Fig. 6). Expression of Scc1RR-DD suppressed premature sister-chromatid separation in *scc1-73* mutant cells but not in *smc3-42* cells (Fig. 6). Thus, Scc1RR-DD alone fulfils the cohesion function of Scc1. The cohesion due to Scc1RR-DD depends on Smc3, as does that produced by wild-type Scc1,

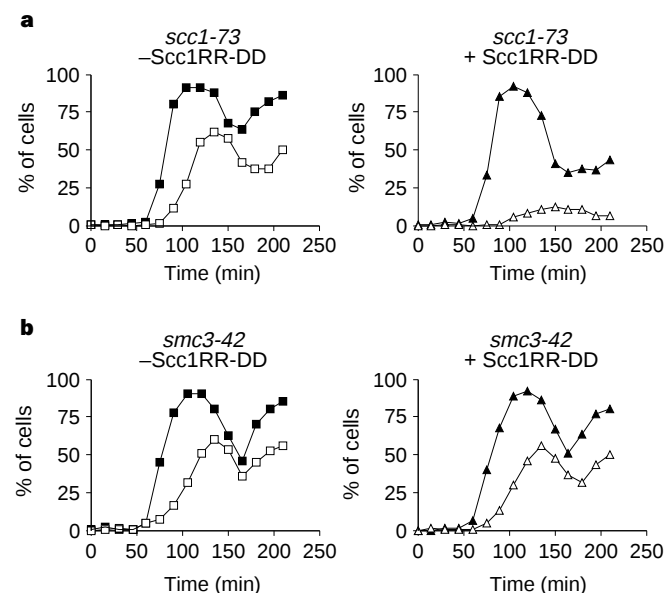


Figure 6 Scc1RR-DD is a functional Scc1 variant. **a**, G1 cells of strain K8103 (*MATa scc1-73 GAL-SCC1(R180D,R268D)-HA3 TetR-GFP TetOs*) were released at 35 °C into medium containing or lacking galactose. The budding index (filled symbols) and the percentage of cells containing separated sister chromatids (open symbols) are shown. **b**, As **a**, except that strain K8149 (*MATa smc3-42 SCC1-myc18 GAL-SCC1(R180D,R268D)-HA3 TetR-GFP TetOs*) was used.

suggesting that Scc1RR-DD differs from wild-type only in its susceptibility to cleavage by Esp1.

Cohesin cleavage triggers anaphase

Our results indicate how cohesion between sister chromatids might be destroyed at the onset of anaphase in yeast (Fig. 7a). As Scc1 is required to maintain cohesion until the metaphase-to-anaphase transition⁷⁻⁹, its disappearance from chromosomes should destroy cohesion. We have shown that cleavage of Scc1 mediated by the separin Esp1 at the exact point when sister chromatids separate is necessary for sister-chromatid separation, for dissociation of Scc1 from chromosomes, and for the movement of spindle poles to opposite ends of the cell. We propose that sister separation in yeast is triggered by the sudden cleavage of Scc1, although what brings this about is not fully understood. Proteolysis of Pds1 shortly before the metaphase-to-anaphase transition is important: Pds1 inhibits Esp1-dependent cleavage *in vitro* and is destroyed suddenly by APC^{Cdc20} *in vivo* so that sister chromatids can separate^{16,17}. The inhibition of APC^{Cdc20} by Mad2 that occurs when chromosomes fail to attach correctly to the mitotic spindle^{18,23,24} prevents the destruction of Pds1 and so blocks the onset of anaphase (hence the dubbing of Pds1 and homologues like Cut2 (ref. 25) as securins). Although destruction of Pds1 is necessary for sister separation, it is not sufficient. The disappearance of Scc1 from chromosomes is tightly regulated in yeast cells lacking Pds1 (ref. 26), so a second pathway must exist that regulates either Esp1 activity or the susceptibility of Scc1 to Esp1-dependent cleavage.

Given that Smc1 and Smc3 may form an antiparallel heterodimer²⁷ and that the link between sisters is a symmetrical, we suggest that Scc1 and Scc3 hold the two Smc1/3 heterodimers together, with each being bound to sister DNA molecules²⁸ (Fig. 7a); separin could then cleave sister chromatids apart. Cleavage also destabilizes

the association of Scc1 with the rest of the cohesin complex. Our results do not indicate whether it is cleavage of Scc1, its subsequent dissociation from the cohesin complex, or both combined that triggers sister-chromatid separation.

Is cohesin cleavage general?

To investigate whether Scc1 was separin's only target, we searched for yeast proteins containing the Scc1 cleavage-site consensus sequences SxExGRR. Only one protein gave a convincing match: we call this protein Rec8 (ref. 29) on the basis of its homology to the *rec8* gene product of fission yeast (Fig. 7b). Rec8 is a member of the Scc1 family and contains two SxExGRK motifs; it is only expressed in meiotic cells and, unlike Scc1, is essential for sister-chromatid cohesion during meiosis I and II (ref. 30; and F. Klein, S. Buonomo and K.N., unpublished results). Rec8 seems to replace Scc1 during meiosis, and cleavage of Rec8 might be crucial for separating sister chromatids during meiosis.

We have not been able to detect Scc1 cleavage motifs in homologous proteins from animals, but the equivalent protein in fission yeast, Rad21 (ref. 31), does contain two near matches in a similar region of the protein to those from Scc1 (Fig. 7b). Most cohesin in animal cells dissociates from chromosomes during prometaphase². A key question for the future is therefore whether the target for animal cell separins is a small fraction of the cohesin pool that persists on metaphase chromosomes or some other cohesion protein.

Is the Esp1 separin a protease?

Although we have not directly determined the identity of the protease that cleaves Scc1, our finding that the cleavage activity in extracts is roughly proportional to their Esp1 concentration (data not shown) raises the possibility that Esp1 is itself the protease. Esp1 and its homologues in other organisms are all large proteins of

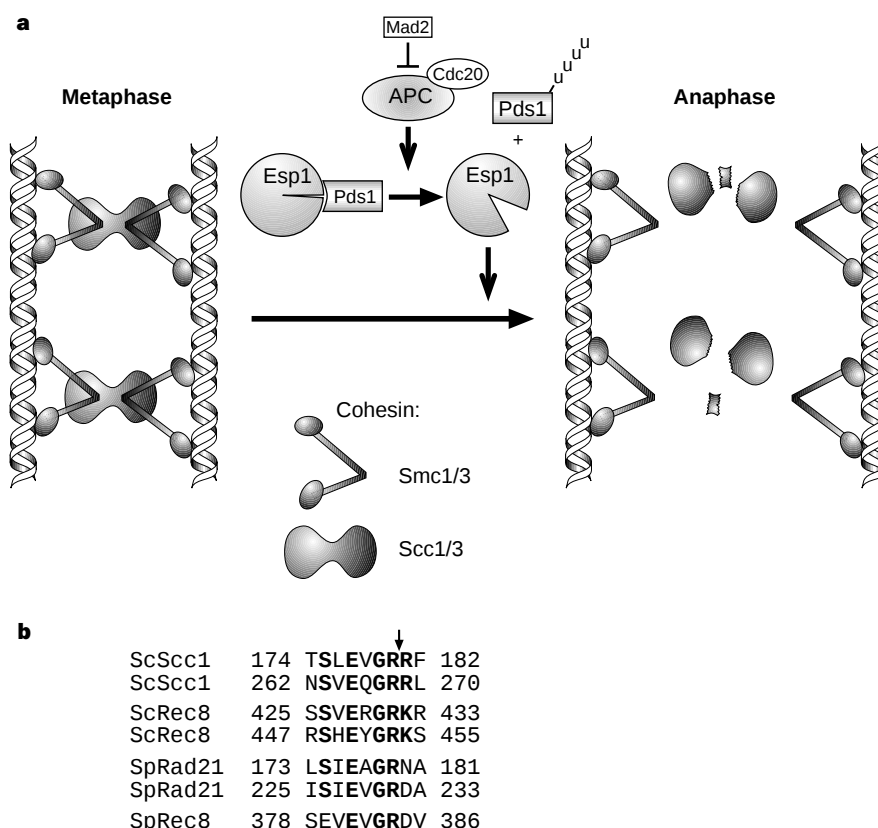


Figure 7 Model for separin action on cohesin, and conservation of potential cohesin cleavage sites. **a**, Proteolytic cleavage of one of cohesin's subunits is necessary for sister separation, suggesting that cohesin complexes link sister chromatids. Mad2 is included in the scheme as a known inhibitor of APC^{Cdc20}

(ref. 24). **b**, Alignment of known and putative cohesin cleavage sites. Sequence motifs similar to the consensus between the two Scc1 cleavage sites were searched for using the HMMER algorithm³⁶.

relative molecular mass ~200K (refs 13, 15, 22, 32). Most of their sequences are not highly conserved, but they all contain a conserved C-terminal domain, which might be responsible for a proteolytic activity. This 'separin domain' does not resemble any known protease. However, the insensitivity of the cleavage reaction to many known protease inhibitors suggests that the protease may have a novel mechanism of action. If Esp1 is not itself the protease, then it might instead be an allosteric effector of a protease, which either resides on chromatin or is present in the soluble fraction. Indeed, Esp1 might have functions in addition to Scc1 cleavage³⁷. It will be necessary to purify Esp1 and provide it with a more clearly defined substrate to answer these questions. □

Methods

Plasmids and strains. The Scc1 coding sequence was cloned under control of the *GAL1-10* promoter into a Ylplac128 derived vector³³. A DNA fragment encoding 3 tandem HA epitopes was inserted into a *NotI* restriction site introduced by PCR at the C terminus of *SCC1*. Site-directed mutagenesis was performed by exchanging restriction fragments from Scc1 with PCR fragments obtained using primers containing the desired nucleotide changes.

All strains used were derivatives of W303. Epitope tags at the C terminus of the endogenous Scc1p were generated by a PCR one-step tagging method (W. Zachariae and K.N., unpublished results). The Myc-epitope tag at the N terminus of endogenous Scc1 was obtained by integration of a N-terminally tagged portion of Scc1 at the *SCC1* locus. Strains expressing Scc1-Myc18, Esp1, and Cdc20 under control of the *GAL1-10* promoter have been described^{7,12,21}.

Cell-growth and cell-cycle experiments. Cells were grown in complete medium³⁴ at 25 °C unless otherwise stated. Strains expressing Cdc20, Esp1 or Scc1 from the *GAL1-10* promoter were grown in complete medium containing 2% raffinose as carbon source. The *GAL1-10* promoter was induced by adding 2% galactose. A G1-like arrest was achieved by adding 1 µg ml⁻¹ of the pheromone α -factor to the medium. For metaphase arrest, 15 µg ml⁻¹ nocodazole was added with 1% DMSO. Metaphase arrest due to Cdc20 depletion was obtained in cells with Cdc20 under control of the *GAL1-10* promoter by shifting to medium containing raffinose only. For release from the arrest, 2% galactose was added back to the culture.

In vitro assay for Esp1 activity. A crude Triton X-100-insoluble chromatin preparation was obtained from yeast cells as described¹⁹. The pelleted chromatin was resuspended in yeast cell extracts that had been prepared similarly to the supernatant fraction of the chromatin preparation. Routinely, one tenth volume of an ATP regenerating system (50 mM HEPES/KOH, pH 7.5, 100 mM KCl, 10 mM MgCl₂, 10 mM ATP, 600 mM creatine phosphate, 1.5 mg ml⁻¹ phosphocreatine kinase, 1 mM DTT, 10% glycerol) was added. (This addition was later found not to be essential.) Reactions were incubated for 10 min at 25 °C with gentle shaking, and stopped on ice. The chromatin fraction was separated again from the supernatant by centrifugation, and resuspended in buffer EBX¹⁹. Equivalent aliquots of supernatant and chromatin pellet were analysed by SDS-PAGE and western blotting. HA-tagged Scc1 was detected using monoclonal antibody 16B12, Myc-tagged Scc1 with monoclonal antibody 9E10. As overexpression of Esp1 from the *GAL1-10* promoter is toxic to cells, extracts with overproduced Esp1 were prepared 2 h after induction with galactose of a culture pregrown in medium containing raffinose only.

Protein sequencing of the Scc1 cleavage site. The C-terminal Scc1 cleavage fragment was isolated from strain K7756 (*MATa cdc20Δ GAL-CDC20 SCC1-myc18*). Cells synchronized in anaphase were obtained as described for Fig. 3. Protein extract of 5 × 10⁹ cells was prepared by breakage with glass beads in breakage buffer (50 mM HEPES/KOH, pH 7.5, 100 mM KCl, 2.5 mM MgCl₂, 1 mM DTT, 0.25% Triton X-100, 0.1% SDS, plus protease inhibitors). Myc-epitope-tagged protein was immunoprecipitated with 20 µg anti-Myc 9E11 monoclonal antibody, resolved on SDS-PAGE and transferred to a PVDF membrane³⁵. N-terminal sequencing of the band corresponding to the Scc1 cleavage fragment was performed using an Applied Biosystems 492A sequencer. The amino-acid sequence was RLGESIM, corresponding to the Scc1 amino-acid residues from position 269.

Chromosome spreading. Analysis of DNA content and chromosome spreading have been described⁷, but spheroplastation to prepare cells for spreading was at 37 °C for only 5 min.

Received 15 March; accepted 12 May 1999.

- Gimenez-Abian, J. F., Clarke, D. J., Mullinger, A. M., Downes, C. S. & Johnson, R. T. A postprophase topoisomerase II-dependent chromatid core separation step in the formation of metaphase chromosomes. *J. Cell. Biol.* **131**, 7–17 (1995).
- Losada, A., Hirano, M. & Hirano, T. Identification of *Xenopus* SMC protein complexes required for sister chromatid cohesion. *Genes Dev.* **12**, 1003–1012 (1998).
- Nicklas, R. B. The forces that move chromosomes in mitosis. *Annu. Rev. Biophys. Chem.* **17**, 431–449 (1988).
- Miyazaki, W. Y. & Orr-Weaver, T. L. Sister chromatid cohesion in mitosis and meiosis. *Annu. Rev. Genet.* **28**, 167–187 (1994).
- Rieder, C. L. & Salmon, E. D. The vertebrate cell kinetochore and its roles during mitosis. *Trends Cell Biol.* **8**, 310–318 (1998).
- Nasmyth, K. Separating sister chromatids. *Trends Biochem. Sci.* **24**, 98–104 (1999).
- Michaelis, C., Ciosk, R. & Nasmyth, K. Cohesins: Chromosomal proteins that prevent premature separation of sister chromatids. *Cell* **91**, 35–45 (1997).
- Guacci, V., Koshland, D. & Strunnikov, A. A direct link between sister chromatid cohesion and chromosome condensation revealed through the analysis of *MCD1* in *S. cerevisiae*. *Cell* **91**, 47–57 (1997).
- Toth, A. *et al.* Yeast cohesin complex requires a conserved protein, Eco1p (Ctf7), to establish cohesion between sister chromatids during DNA replication. *Genes Dev.* **13**, 320–333 (1999).
- Uhlmann, F. & Nasmyth, K. Cohesion between sister chromatids must be established during DNA replication. *Curr. Biol.* **8**, 1095–1101 (1998).
- Skibbens, R. V., Corson, L. B., Koshland, D. & Hieter, P. Ctf7p is essential for sister chromatid cohesion and links mitotic chromosome structure to the DNA replication machinery. *Genes Dev.* **13**, 307–319 (1999).
- Ciosk, R. *et al.* An ESP1/PDS1 complex regulates loss of sister chromatid cohesion at the metaphase to anaphase transition in yeast. *Cell* **93**, 1067–1076 (1998).
- May, G. S., McGoldrick, C. A., Holt, C. L. & Denison, S. H. The *bimB3* mutation of *Aspergillus nidulans* uncouples DNA replication from the completion of mitosis. *J. Biol. Chem.* **267**, 15737–15743 (1992).
- Funabiki, H., Kumada, K. & Yanagida, M. Fission yeast Cut1 and Cut2 are essential for sister chromatid separation, concentrate along the metaphase spindle and form large complexes. *EMBO J.* **15**, 6617–6628 (1996).
- Nagase, T., Seki, N., Ishikawa, K., Tanaka, A. & Nomura, N. Prediction of the coding sequences of unidentified human genes. V. The coding sequences of 40 new genes (K1A0161–K1A0200) deduced by analysis of cDNA clones from human cell line KG-1. *DNA Res.* **3**, 17–24 (1996).
- Cohen-Fix, O., Peters, J.-M., Kirschner, M. W. & Koshland, D. Anaphase initiation in *Saccharomyces cerevisiae* is controlled by the APC-dependent degradation of the anaphase inhibitor Pds1p. *Genes Dev.* **10**, 3081–3093 (1996).
- Peters, J.-M. SCF and APC the Yin and Yang of cell cycle regulated proteolysis. *Curr. Opin. Cell Biol.* **10**, 759–768 (1998).
- Yamamoto, A., Guacci, V. & Koshland, D. Pds1p, an inhibitor of anaphase in budding yeast, plays a critical role in the APC and checkpoint pathway(s). *J. Cell Biol.* **133**, 99–110 (1996).
- Liang, C. & Stillman, B. Persistent initiation of DNA replication and chromatin-bound MCM proteins during the cell cycle in *cdc6* mutants. *Genes Dev.* **11**, 3375–3386 (1997).
- Groigno, L. & Whitaker, M. An anaphase calcium signal controls chromosome disjunction in urchin embryos. *Cell* **92**, 193–204 (1998).
- Lim, H. H., Goh, P.-Y. & Surana, U. Cdc20 is essential for the cyclosome-mediated proteolysis of both Pds1 and Clb2 during M phase in budding yeast. *Curr. Biol.* **8**, 231–234 (1998).
- McGrew, J. T., Goetsch, L., Byers, B. & Baum, P. Requirement for ESP1 in the nuclear division of *S. cerevisiae*. *Mol. Biol. Cell* **3**, 1443–1454 (1992).
- Rudner, A. D. & Murray, A. W. The spindle assembly checkpoint. *Curr. Opin. Cell Biol.* **8**, 773–780 (1996).
- Amon, A. The spindle checkpoint. *Curr. Opin. Gen. Dev.* **9**, 69–75 (1999).
- Funabiki, H. *et al.* Cut2 proteolysis required for sister-chromatid separation in fission yeast. *Nature* **381**, 438–441 (1996).
- Alexandru, G., Zachariae, W., Schleiffer, A. & Nasmyth, K. Sister chromatid separation and chromosome re-duplication are regulated by different mechanisms in response to spindle damage. *EMBO J.* **18**, 2707–2721 (1999).
- Melby, T. E., Ciampaglio, C. N., Briscoe, G. & Erickson, H. P. The symmetrical structure of structural maintenance of chromosomes (SMC) and MukB proteins: long, antiparallel coiled coils, folded at a flexible hinge. *J. Cell Biol.* **142**, 1595–1604 (1998).
- Akhmedov, A. T. *et al.* Structural maintenance of chromosomes protein C-terminal domains bind preferentially to DNA with secondary structure. *J. Biol. Chem.* **273**, 24088–24093 (1998).
- Lin, Y., Larson, K. L., Dorer, R. & Smith, G. R. Meiotically induced *rec7* and *rec8* genes of *Schizosaccharomyces pombe*. *Genetics* **132**, 75–85 (1992).
- Molnar, M., Bahler, J., Sipicki, M. & Kohli, J. The *rec8* gene of *Schizosaccharomyces pombe* is involved in linear element formation, chromosome pairing and sister-chromatid cohesion during meiosis. *Genetics* **141**, 61–73 (1995).
- Birkenbihl, R. P. & Subramani, S. The *rad21* gene product of *Schizosaccharomyces pombe* is a nuclear, cell cycle-regulated phosphoprotein. *J. Biol. Chem.* **270**, 7703–7711 (1995).
- Uzawa, S., Samejima, I., Hirano, T., Tanaka, K. & Yanagida, M. The fission yeast cut1+ gene regulates spindle pole duplication and has homology to the budding yeast ESP1 gene. *Cell* **62**, 913–925 (1990).
- Gietz, R. D. & Sugino, A. New yeast–*Escherichia coli* shuttle vectors constructed with *in vitro* mutagenized yeast genes lacking six base pair restriction sites. *Gene* **74**, 527–534 (1988).
- Rose, M. D., Winston, F. & Hieter, P. *Laboratory Course Manual for Methods in Yeast Genetics* (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, 1990).
- Eckerskorn, C. & Lottspeich, F. Structural characterization of blotting membranes and the influence of membrane parameters for electroblotting and subsequent amino acid sequence analysis of proteins. *Electrophoresis* **14**, 831–838 (1993).
- Eddy, S. R. Hidden Markov models. *Curr. Opin. Struct. Biol.* **6**, 361–365 (1996).
- Kumada, K. *et al.* Cut1 is loaded onto the spindle by binding to Cut2 and promotes anaphase spindle movement upon Cut2 proteolysis. *Curr. Biol.* **8**, 633–641 (1998).

Acknowledgements. We thank E. Kramer for *in vitro* translated Pds1; K. Mechtler for antibody purification; L. Huber for rabbit anti-Myc antibody; T. Skern for discussion; R. Ciosk, C. Michaelis, A. Toth and W. Zachariae for discussion and for yeast strains; A. Schleiffer for database searches; H. Tkadletz for graphics; and M. Grotzer and J.-M. Peters for comments on the manuscript. F.U. acknowledges support from an EMBO long-term fellowship.

Correspondence and requests for materials should be addressed to K.N. (e-mail: nasmyth@nt.imp.univie.ac.at).

Spin gap and magnetic coherence in a clean high-temperature superconductor

B. Lake*†‡, G. Aeppli§, T. E. Mason*†, A. Schröder||, D. F. McMorrow‡, K. Lefmann‡, M. Isshiki¶, M. Nohara¶, H. Takagi¶ & S. M. Hayden#

* Department of Physics, University of Toronto, Toronto, ON M5S 1A7, Canada

† Oak Ridge National Laboratory, Oak Ridge, Tennessee 37831, USA

‡ Department of Condensed Matter Physics and Chemistry,

Risø National Laboratory, 4000 Roskilde, Denmark

§ NEC Research, 4 Independence Way, Princeton, New Jersey 08540, USA

|| Department of Physics, University of Karlsruhe, D-76128 Karlsruhe, Germany

¶ Institute for Solid State Physics, University of Tokyo, Roppongi 7-22-1,

Minato-ku, Tokyo 106-8666, Japan

H. H. Wills Physics Laboratory, University of Bristol, Bristol BS8 1TL, UK

A notable aspect of high-temperature superconductivity in the copper oxides is the unconventional nature of the underlying paired-electron state. A direct manifestation of the unconventional state is a pairing energy—that is, the energy required to remove one electron from the superconductor—that varies (between zero and a maximum value) as a function of momentum, or wavevector^{1,2}: the pairing energy for conventional superconductors is wavevector-independent^{3,4}. The wavefunction describing the superconducting state will include the pairing not only of charges, but also of the spins of the paired charges. Each pair is usually in the form of a spin singlet⁵, so there will also be a pairing energy associated with transforming the spin singlet into the higher-energy spin triplet form without necessarily unbinding the charges. Here we use inelastic neutron scattering to determine the wavevector-dependence of spin pairing in $\text{La}_{2-x}\text{Sr}_x\text{CuO}_4$, the simplest high-temperature superconductor. We find that the spin pairing energy (or 'spin gap') is wavevector independent, even though superconductivity significantly alters the wavevector dependence of the spin fluctuations at higher energies.

The experimental technique that we use is inelastic neutron scattering, for which the cross-section is directly proportional to the magnetic excitation spectrum and can be used to probe it as a function of wavevector and energy transfer. In addition, we have selected $\text{La}_{2-x}\text{Sr}_x\text{CuO}_4$, the simplest of the high-temperature (high- T_c) superconductors. The materials consist of nearly square CuO_2 lattices with Cu atoms at the vertices and O atoms on the edges alternating with LaSrO charge reservoir layers. In the absence of Sr doping, the compound is an antiferromagnetic insulator, where the spin on each Cu^{2+} ion is antiparallel to those on its four nearest neighbours. Because of the unit cell doubling, magnetic Bragg reflections appear at wavevectors such as $(\frac{1}{2}, \frac{1}{2})$ (sometimes called (π, π) in the two-dimensional reciprocal space of the CuO_2 planes⁶. Doping yields a superconductor without long-range magnetic order but which has low-energy magnetic excitations peaked at the quartet of wavevectors $\mathbf{Q}_\delta = (\frac{1}{2}(1 \pm \delta), \frac{1}{2})$ and $(\frac{1}{2}, \frac{1}{2}(1 \pm \delta))$, shown in Fig. 1a. The recent discovery of nearly identical fluctuations in the high- T_c $\text{YBa}_2\text{Cu}_3\text{O}_{7-y}$ bilayer materials⁷ clearly indicates their relevance to the larger issue of high- T_c superconductivity, and validates the continued study of $\text{La}_{2-x}\text{Sr}_x\text{CuO}_4$ as the copper oxide with the least structural and electronic complexity.

The samples are single-crystal rods grown in an optical image furnace. The most reliable measure of the quality of bulk superconductors is the specific heat C . For our samples, there is a jump of $\Delta C/k_B T_c = 7 \text{ mJ mol}^{-1}$ at $T_c = 38.5 \text{ K}$. As $T \rightarrow 0$, $C = \gamma_s T$ where γ_s is proportional to the electronic density of states at the Fermi level

and has the value $\gamma_s < 0.8 \text{ mJ mol}^{-1} \text{ K}^{-2}$. This, together with an estimate of $10 \text{ mJ mol}^{-1} \text{ K}^{-2}$ for the corresponding normal state γ_N , indicates that the bulk superconducting volume fraction $1 - \gamma_s/\gamma_N$ of our samples is greater than 0.9. This, as well as the high value of T_c and the narrowness of the transition, is evidence for the very high quality of our large, single crystals. The basic experimental configurations are similar to those employed previously^{8,9}. Figure 1a shows the reciprocal space regions probed. A series of scans like those indicated in the figure, performed for a range of energy transfers, were used to build up the $\mathbf{Q}$ - E maps in Fig. 1b and c, which show the scattering around the incommensurate peaks in the normal and superconducting states (here E is the energy transfer).

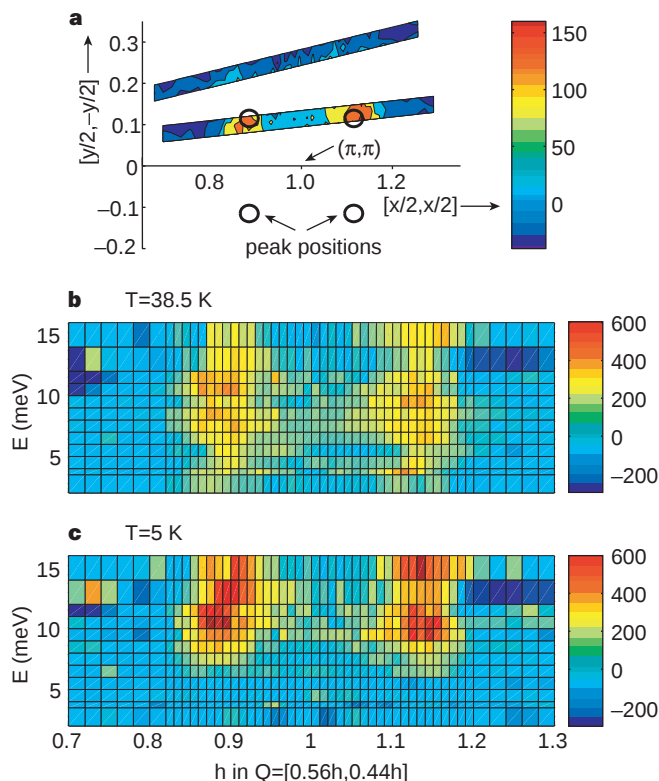


Figure 1 The reciprocal-space regions over which measurements were made and the resulting data as a function of wavevector and energy transfer. **a**, Reciprocal-space diagram of the CuO_2 planes in $\text{La}_{2-x}\text{Sr}_x\text{CuO}_4$ where the black circles represent the incommensurate peaks surrounding (π, π) . The coloured strips give the data collected at an energy transfer of 9 meV and temperature of 5 K, and show the areas probed in this experiment. The lower strip passes through two of the peaks and provides the signal, whereas the upper strip lies far from the peaks and is used as the background. **b**, A $\mathbf{Q}$ - E map measured at the transition temperature of 38.5 K, for the wavevectors shown in **a** and energy transfers from 2 to 16 meV; the parameter h defines the two-dimensional wavevector $\mathbf{Q} = [0.56h, 0.44h]$. The data displayed are background-subtracted and the thermal population factor $(\exp(-E/k_B T) - 1)^{-1}$ has been divided out to give the quantity $S(\mathbf{a}, \omega)$. The colouring of the squares indicates the intensity observed in units of counts per 15 min. **c**, A similar map to **b** but for the superconducting phase at 5 K. The sample used for these measurements consisted of five crystals grown by the travelling solvent floating zone method; each was $\sim 4 \text{ mm}$ in diameter and 20 mm long. To maximize signal, the crystals were mounted on a single holder so that their axes were parallel to within 0.8° . The measurements were performed on the RITA spectrometer at Risø National Laboratory³⁰. RITA differs from its predecessor, TAS6, by use of a velocity selector for better filtering of the incident beam, superior beam optics between the reactor and the sample, and a large position-sensitive detector. In this experiment pixels at different vertical heights on the detector were binned separately so that a single instrumental scan produced a two-dimensional plot in reciprocal space as seen in **a**.

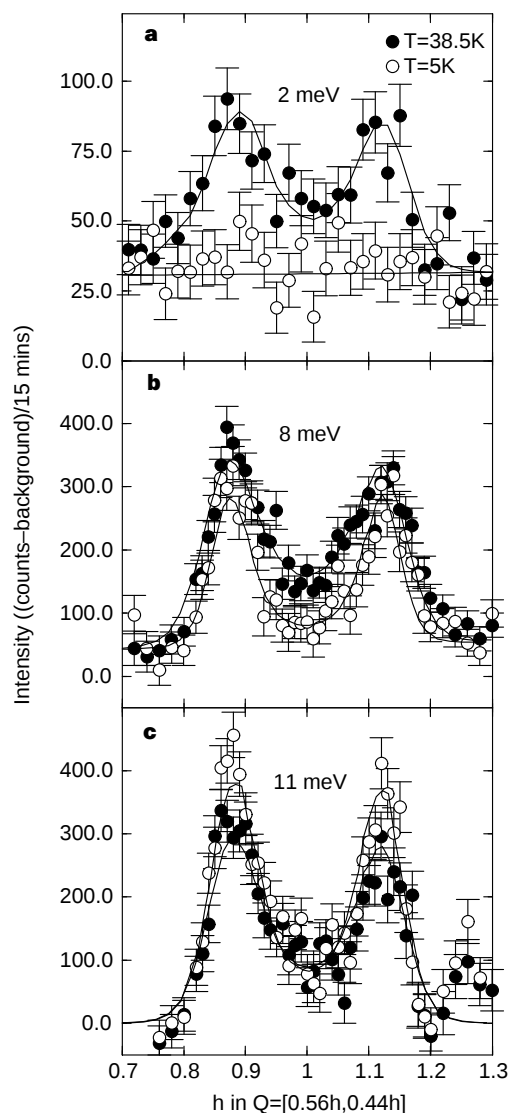


Figure 2 Constant- E scans through the incommensurate peaks at T_c and in the superconducting phase at 5 K. **a**, The data for 2 meV well below the spin-gap energy of 6.7 meV. The incommensurate peaks are clearly visible at T_c but are completely suppressed in the superconducting phase. **b**, The scattering at 8 meV, somewhat above the gap energy. The peak amplitudes are approximately equal at the two temperatures, while their lineshapes are different with scattering at 5 K narrower than at 38.5 K. **c**, The scattering at 11 meV, well above the spin gap. The scattering amplitude at 5 K is significantly greater than at 38.5 K. The solid lines through the data were obtained by following the analysis procedures of ref. 10.

Figure 1b shows that the normal-state excitations at 38.5 K are localized near Q_δ but are entirely delocalized in E . In other words, the magnetic fluctuations which are favoured are those with a particular spatial period $1/\delta$ corresponding to Q_δ , but no particular temporal period. Cooling below T_c produces a very different image in Q - E space. In Fig. 1c all low-frequency excitations ($E \leq 5$ meV) seem to be eliminated, and there is an enhancement of the signal above 8 meV at the incommensurate wavevectors. The signal now has obvious peaks at around $E = 11$ meV and $\delta = 0.29 \pm 0.03$ reciprocal lattice units (r.l.u.). We can thus visualize the zero point fluctuations in the superconductor as magnetic density waves undergoing (damped) oscillation with a frequency of 2.75 THz. In the normal paramagnetic state, the motion of the density waves becomes entirely incoherent.

Figure 2a-c show a series of constant- E sections through the data in Fig. 1. These graphs show that superconductivity induces a

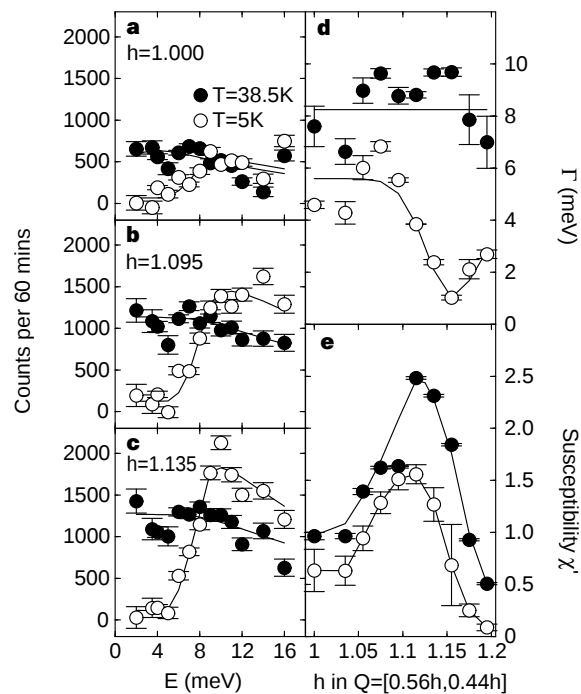


Figure 3 Spectra at various wavevectors, and the Q -dependence of the inverse lifetime and susceptibility extracted by fitting such profiles. (The wavevectors range from between the incommensurate peaks up to the peak maxima.) Panel **a** shows the constant- Q spectrum at $h = 1.000$ (a wavevector exactly in between the two peaks), **b** shows the spectrum at $h = 1.095$ (a position on the inner slope of the peak) and **c** gives the spectrum at $h = 1.135$ (the peak maximum). The spin gap is present at all three wavevectors in the superconducting phase with the scattering eliminated below 6 meV. As h is changed from 1.000 to 1.135 there is a clear maximum in the response at energies around 11 meV in the superconducting phase. Panels **d** and **e** show the values of the energy scale Γ and the real part of the magnetic susceptibility χ' (in units of counts per 15 min) as functions of wavevector for both 38.5 K and 5 K. These quantities were extracted by fitting to the data the lineshape given in equation (1) convolved with the instrumental resolution. Γ is approximately constant at T_c but strongly Q -dependent at 5 K and χ' is smaller at 5 K than at 38.5 K although it has a similar wavevector-dependent profile at the two temperatures.

complete loss of signal for $E = 2$ meV (Fig. 2a), a significant intensity-preserving sharpening of the incommensurate peaks for $E = 8$ meV (Fig. 2b), and a large enhancement of the peaks for $E = 11$ meV (Fig. 2c). The peak narrowing in Fig. 2b and c corresponds to a spectacular superconductivity-induced rise in the magnetic coherence lengths (defined as the resolution-corrected inverse half-widths at half-maxima obtained as in ref. 10) from 20.1 ± 0.9 to 33.5 ± 2.0 Å, and from 25.5 ± 0.1 to 34.3 ± 0.8 Å, respectively.

Figure 3a-c displays constant- Q spectra both away from Q_δ (Fig. 3a and b) and at Q_δ (Fig. 3c). Superconductivity removes the low- E signal below a threshold energy, while it enhances the higher- E signal close to Q_δ . The threshold for $T < T_c$ appears the same for the three wavevectors shown in Fig. 3a-c, with the increase in intensity first visible in all cases at 6 meV. To quantify how superconductivity changes the spectra, we fit the data with the convolution of the instrumental resolution (full-width at half-maximum, 2 meV) and

$$S(Q, E) = \frac{AE'\Gamma}{(\Gamma^2 + E^2)(1 - \exp(E/k_B T))} \quad (1)$$

where

$$E' = \text{Re}\{[(E - \Delta + i\Gamma_s)(E + \Delta + i\Gamma_s)]^{1/2}\} \quad (2)$$

and A is the amplitude, Δ is the spin gap, Γ is the inverse lifetime of spin fluctuations with $E \gg \Delta$ (if $\Delta \ll \Gamma$), E' is an odd function of E

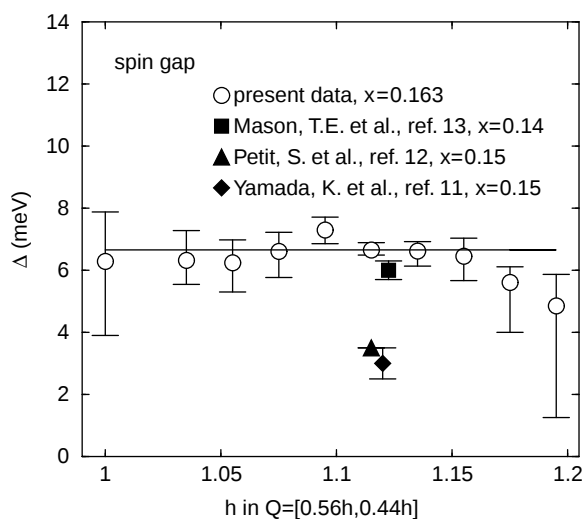


Figure 4 The wavevector dependence of the spin gap in the superconducting state at 5 K. The gap was extracted by fitting equation (1), convolved with the instrumental resolution, to constant- Q spectra such as those shown in Fig. 3a–c. The gap values are given by the open circles and are independent of wavevector. The solid line gives the fitted wavevector-independent value of the gap which is $\Delta = 6.7$ meV. The solid symbols show the spin gaps determined previously for $\text{La}_{2-x}\text{Sr}_x\text{CuO}_4$. The filled square gives the gap for an $x = 0.14$ sample; the gap was extracted by the same method as used here and found to be of a similar size. The filled triangle and the filled diamond give the gap for two different $x = 0.15$ samples; in both cases the gap was defined as the threshold for visible scattering without considering resolution broadening or damping at the gap edge. Using this definition the gap was found to have a much smaller value. In all except the present work, the spin gap was established at only a single wavevector.

which defines the degree to which the spectrum has a gap, and Γ_s is the inverse lifetime of the fluctuations at the gap edge.

In the normal state, the best fits are obtained for $\Delta = 0$ meV, and the fitted value of Γ is essentially Q -independent (Fig. 3d). Thus, the lower-amplitude fluctuations with wavevectors different from Q_0 have lifetimes similar to those at the incommensurate peak positions. The Q dependence of the signal is entirely accounted for by the Q dependence of the real part $\chi'(Q)$ of the magnetic susceptibility (Fig. 3e) which, when $\Delta = 0$, is simply the amplitude A . In the superconducting state, Γ (Fig. 3d), which characterizes the shape of the spectrum well above the spin gap, becomes strongly Q dependent. At the same time, $\chi'(Q)$ (Fig. 3e), related via a Kramers–Kronig relation to the parameters in equation (1), is suppressed. This explicitly demonstrates that superconductivity reduces the tendency towards static incommensurate magnetic order in $\text{La}_{2-x}\text{Sr}_x\text{CuO}_4$.

Figure 4 shows the Q dependence of the spin gap Δ . As anticipated from inspection of the data in Fig. 3, Δ is Q -independent and has the value 6.7 meV. The gap is quite sharp for our sample, with $\Gamma_s \leq 0.2$ meV for all Q . Also shown in fig. 4 are the results for $x = 0.15$ (refs 11, 12) and 0.14 (ref. 10). $\Delta(Q_0)$ is indistinguishable for the present $x = 0.163$ and the older $x = 0.14$ samples; the difference in the low- E behaviour is primarily due to the much larger damping ($\Gamma_x = 1.2$ meV for $x = 0.14$; ref. 13). In addition, the Q independence of $\Delta(Q)$ is consistent with the Q -independent but incomplete suppression of the magnetic fluctuations in the $x = 0.14$ sample¹⁴. In contrast, the results of refs 11 and 12 show a large discrepancy with $x = 0.163$, where the spin gap quoted in these papers is defined as the threshold for visible scattering. Nevertheless, the results of ref. 11 are consistent with our work if we use the definition—advocated here and in ref. 13—of Δ given by equation (2). Fitting the data of ref. 11 to equation (1) with $\Delta = 6.7$ meV yields $\Gamma_s = 0.5$ meV, a value intermediate between our findings of 1.2 and 0.1 meV for $x = 0.14$ and $x = 0.163$.

Our experiments show that superconductivity produces strongly momentum-dependent changes in the magnetic excitations with energies above a momentum-independent spin gap. The data in their entirety do not resemble the predictions^{15–20} for any superconductors, be they s -wave or d -wave. Most notably, all d -wave theories anticipate dispersion in the spin gap which would have been observed over the wavevector range and for the energy resolution of the present experiment. At the same time, s -wave theory cannot account for the value of the spin gap. We are unaware of calculations which yield the dramatic incommensurate peak sharpening and enhancements that we see above the spin gap while at the same time showing a large reduction in the real part of the magnetic susceptibility.

There are other difficulties with the conventional weak-coupling d -wave approach, which posits nodes and therefore a smaller relative superconductivity-induced reduction in scattering between, rather than at, the incommensurate peaks. Figure 2b shows the opposite—just above the gap energy, the incommensurate peak intensities are preserved while the scattering between the peaks is suppressed. Furthermore, the peak sharpening in momentum space for $\hbar\omega > \Delta$ finds a precedence only in quantum systems, such as $S = 1$ antiferromagnetic (Haldane) spin chains and rotons in superfluid helium, which have well defined gaps with non-zero minima. Thus, although our statistics and resolution cannot exclude a small population of spin-carrying subgap quasiparticles, the systematics of the signal found near the gap energy make such quasiparticles improbable.

As for any other spectroscopic experiment, we can only place an upper bound on the signal below the dispersionless gap. Inspection of Fig. 3 shows that in between the incommensurate peaks at $Q = (\frac{1}{2}(1 + \frac{\delta}{2}), \frac{1}{2}(1 - \frac{\delta}{2}))$, where ordinary weak-coupling d -wave theories generally anticipate nodes in the spin gap, the intensity for 2 meV at 5 K is less than 14% of what was seen at T_c and below 5% of that observed for the incommensurate peaks at 5 K.

Given the overwhelming evidence for d -wave superconductivity in the hole-doped high- T_c superconductors^{1,2,21,22}, we see our data not as evidence against d -wave superconductivity but as proof that the spin excitations in the superconducting state do not parallel the charge excitations in the fashion assumed for ordinary d - and s -wave superconductors. Our measurements, which are sensitive exclusively to the spin sector, taken together with the evidence for d -wave superconductivity in the charge sector suggest that the high- T_c superconductors are in fact Luther–Emery (LE) liquids: that is, they are materials with gapped (triplet) spin excitations and gapless spin-zero charge excitations^{23,24}. Luther–Emery liquids arise in one-dimensional interacting Fermi systems, which formally resemble two-dimensional d -wave superconductors—the dimensionality (zero) of the nodal points where the gap vanishes in the two-dimensional copper oxide is the same as that of the Fermi surface of a one-dimensional metal. There are other arguments for the applicability of the concept of Luther–Emery liquids. The first is that theory indicates that such liquids are the ground states of ladder compounds, one-dimensional strips of finite width cut from CuO_2 planes^{25–28}. The second involves the breakdown of spin–charge separation when the spin gap collapses to zero, which can be brought about by a magnetic field whose Zeeman energy matches the spin-gap energy. The 6.7-meV spin gap which we measure is much closer to the Zeeman energy of the upper critical field measured²⁹ for samples similar to ours than to an ordinary Bardeen–Cooper–Schrieffer pairing energy $\geq 3.5k_B T_c = 11.6$ meV. □

Received 15 April; accepted 4 May 1999.

- Shen, Z.-X. *et al.* Anomalously large gap anisotropy in the a – b plane of $\text{Bi}_2\text{Sr}_2\text{CaCu}_2\text{O}_{8+y}$. *Phys. Rev. Lett.* **70**, 1553–1556 (1993).
- Ding, H. *et al.* Angle-resolved photoemission spectroscopy study of the superconducting gap anisotropy in $\text{Bi}_2\text{Sr}_2\text{CaCu}_2\text{O}_{8+y}$. *Phys. Rev. B* **54**, R9678–R9681 (1996).
- Miyake, K., Schmitt-Rink, S. & Varma, C. M. Spin-fluctuation-mediated even-parity pairing in heavy-fermion superconductors. *Phys. Rev. B* **34**, 6554–6556 (1986).

4. Scalapino, D. J., Loh, E. Jr & Hirsch, J. E. *d*-wave pairing near a spin-density-wave instability. *Phys. Rev. B* **34**, 8190–8195 (1986).
5. Bardeen, J., Cooper, L. N. & Schrieffer, J. R. Theory of superconductivity. *Phys. Rev.* **61**, 1175–1185 (1957).
6. Vaknin, D. *et al.* Antiferromagnetism in La_2CuO_4 . *Phys. Rev. Lett.* **58**, 2802–2805 (1987).
7. Mook, H. A. *et al.* Spin fluctuations in $\text{YBa}_2\text{Cu}_3\text{O}_{6.6}$. *Nature* **395**, 580–582 (1998).
8. Cheong, S.-W. Incommensurate magnetic fluctuations in $\text{La}_{2-x}\text{Sr}_x\text{CuO}_4$. *Phys. Rev. Lett.* **67**, 1791–1794 (1991).
9. Mason, T. E., Aeppli, G. & Mook, H. A. Magnetic dynamics of superconducting $\text{La}_{1.86}\text{Sr}_{0.14}\text{CuO}_4$. *Phys. Rev. Lett.* **68**, 1414–1417 (1992).
10. Mason, T. E., Schröder, A., Aeppli, G., Mook, H. A. & Hayden, S. M. New magnetic coherence effect in superconducting $\text{La}_{2-x}\text{Sr}_x\text{CuO}_4$. *Phys. Rev. Lett.* **77**, 1604–1607 (1996).
11. Yamada, K. *et al.* Direct observation of a magnetic gap in superconducting $\text{La}_{1.85}\text{Sr}_{0.15}\text{CuO}_4$ ($T_c = 37.3$ K). *Phys. Rev. Lett.* **75**, 1626–1629 (1995).
12. Petit, S., Moudren, A. H., Hennen, B., Vietkin, A. & Revcolevschi, A. Spin dynamics study of $\text{La}_{2-x}\text{Sr}_x\text{CuO}_4$ by inelastic neutron scattering. *Physica B* **234–236**, 800–802 (1997).
13. Mason, T. E., Aeppli, G., Hayden, S. M., Ramirez, A. P. & Mook, H. A. Magnetic fluctuations in superconducting $\text{La}_{2-x}\text{Sr}_x\text{CuO}_4$. *Physica B* **199–200**, 284–287 (1994).
14. Mason, T. E., Aeppli, G., Hayden, S. M., Ramirez, A. P. & Mook, H. A. Low energy excitations in superconducting $\text{La}_{1.86}\text{Sr}_{0.14}\text{CuO}_4$. *Phys. Rev. Lett.* **71**, 919–921 (1993).
15. Joynt, R. & Rice, T. M. Magnetic properties of anisotropic superconductors. *Phys. Rev. B* **38**, 2345–2353 (1988).
16. Tanamoto, T., Kohno, H. & Fukuyama, H. Magnetic properties of extended *t*-*J* model. II. Dynamical properties. *J. Phys. Soc. Jpn* **63**, 2739–2759 (1994).
17. Rendell, J. M. & Carbotte, J. P. Effects of gap and band anisotropy on spin susceptibility in the oxide superconductors. *Phys. Rev. B* **53**, 589–5899 (1996).
18. Zha, Y., Levin, K. & Si, Q. Neutron experiments as a test of anisotropic pairing in $\text{YBa}_2\text{Cu}_3\text{O}_{7-y}$ and $\text{La}_{2-x}\text{Sr}_x\text{CuO}_4$. *Phys. Rev. B* **47**, 9124–9127 (1993).
19. Bulut, N. & Scalapino, D. J. Neutron scattering in a $d_{x^2-y^2}$ -wave superconductor. *Phys. Rev. B* **50**, 16078–16081 (1993).
20. Morr, D. K. & Pines, D. The effect of superconductivity on the incommensurate magnetic response of cuprate superconductors. Preprint cond-mat/9807214 at (<http://xxx.lanl.gov>) (1998).
21. Tsuei, C. C. *et al.* Pairing symmetry in single-layer tetragonal $\text{Tl}_2\text{Ba}_2\text{CuO}_{6+\delta}$ superconductors. *Science* **271**, 329–332 (1996).
22. Chen, X. K., Irwin, J. C., Troadahl, H. J., Kimura, T. & Kishio, K. Investigation of the superconducting gap $\text{La}_{2-x}\text{Sr}_x\text{CuO}_4$ by Raman spectroscopy. *Phys. Rev. Lett.* **73**, 3290–3292 (1994).
23. Luther, A. & Emery, V. J. Backward scattering in the one-dimensional electron gas. *Phys. Rev. Lett.* **33**, 589–592 (1974).
24. Rokhsar, D. S. & Kivelson, S. A. Superconductivity and the quantum hard-core dimer gas. *Phys. Rev. Lett.* **61**, 2376–2379 (1988).
25. Tsunetsugu, H. Doped spin-liquid antiferromagnets and Luther-Emery liquid. *Physica B* **237–238**, 108–111 (1997).
26. Dagotto, E., Riera, J. & Scalapino, D. Superconductivity in ladders and coupled planes. *Phys. Rev. B* **45**, 5744–5747 (1992).
27. Rice, T. M., Gopalan, S. & Sigrist, M. Superconductivity, spin gaps and Luttinger liquids in a class of cuprates. *Europhys. Lett.* **23**, 445–449 (1993).
28. Emery, V. J., Kivelson, S. A. & Zachar, O. Spin-gap proximity effect mechanism of high-temperature superconductivity. *Phys. Rev. B* **56**, 6120–6147 (1997).
29. Andoh, Y. *et al.* Resistive upper critical fields and irreversibility lines of optimally doped high- T_c cuprates. Preprint (Bell Laboratories, Murray Hill, NJ 07974, USA, 1998).
30. Mason, T. E., Clausen, K. N., Aeppli, G., McMorris, D. F. & Kjems, J. K. RITA: The reinvented triple-axis spectrometer. *Can. J. Phys.* **73**, 697–702 (1995).

Acknowledgements. We thank K. N. Clausen for help and support during the experiments, and B. Batlogg, G. Boebinger, V. Emery, S. Kivelson, H. Mook, D. Morr, D. Pines, Z.-X. Shen, C.-C. Tsuei and J. Zaanen for discussions. Work done at the University of Toronto was sponsored by the Natural Sciences and Engineering Research Council and the Canadian Institute for Advanced Research, while work done at Oak Ridge was supported by the US DOE. T.E.M. was supported by the Alfred P. Sloan Foundation, and A.S. was supported by the TMR program.

Correspondence and requests for materials should be addressed to B.L. (e-mail: bella@phonon.ssd.ornl.gov).

Origin of the Invar effect in iron–nickel alloys

Mark van Schilfgaarde*, I. A. Abrikosov† & B. Johansson†

* Sandia National Laboratories, Livermore, California 94551, USA

† Condensed Matter Theory Group, Physics Department, Uppsala University, S-75121 Uppsala, Sweden

In 1897 Guillaume¹ discovered that face-centred cubic alloys of iron and nickel with a nickel concentration of around 35 atomic per cent exhibit anomalously low (almost zero) thermal expansion over a wide temperature range. This effect, known as the Invar effect, has since been found in various ordered and random alloys and even in amorphous materials². Other physical properties of Invar systems, such as atomic volume, elastic modulus, heat capacity, magnetization and Curie (or Néel) temperature, also show anomalous behaviour. Invar alloys are used in instrumentation, for example as hair springs in watches. It has long been realized that the effect is related to magnetism^{2,3}, but a full understanding is still lacking. Here we present *ab initio* calcula-

tions of the volume dependences of magnetic and thermodynamic properties for the most typical Invar system, a random face-centred cubic iron–nickel alloy, in which we allow for non-collinear spin alignments—that is, spins that may be canted with respect to the average magnetization direction. We find that the magnetic structure is characterized, even at zero temperature, by a continuous transition from the ferromagnetic state at high volumes to a disordered non-collinear configuration at low volumes. There is an additional, comparable contribution to the net magnetization from the changes in the amplitudes of the local magnetic moments. The non-collinearity gives rise to an anomalous volume dependence of the binding energy, and explains other peculiarities of Invar systems.

The discovery of the Invar effect has stimulated many experimental and theoretical studies of this phenomenon, and a number of theoretical models have been suggested for its explanation². The most popular model can be traced back to Weiss³ who introduced the so-called 2 γ -state model. According to this model there are two possible states for face-centred cubic (f.c.c.) γ -Fe: the ferromagnetic high-volume state and the antiferromagnetic low-volume state. Thermal excitations between these two states are supposed to compensate for the usual lattice expansion related to anharmonic effects of the lattice vibrations. Modern electronic-structure theory, improving the basic understanding of this model, seems to confirm its main idea. First-principles calculations of ferromagnetic γ -Fe, ordered Fe_3Ni and random f.c.c. Fe–Ni alloys clearly show the existence of two magnetic states, a low-spin low-volume state (LS) and a high-spin high-volume state (HS)^{2,4–14}. Close to the Invar concentration (~ 35 at. % Ni for the Fe–Ni alloy), the binding energy curves in these systems were found to consist of two distinct branches with a very small energy difference between their respective minima, and this has been viewed as a success of the two-state model².

On the other hand, there are several experimental observations that contradict the 2 γ -state model. For example, the phase transition from the HS state to the LS state as a function of concentration should be of first order^{11,13,14}, while experimentally it is not³. Moreover, calculations overestimate the volume decrease as a function of concentration^{13,15}. Finally, the two-minimum shape of the binding energy curve should lead to some discontinuities in the pressure dependence of physical properties which to our knowledge have not been seen in any experiment².

In the case of random alloys the so-called disordered local moment model was formulated to account for the possibility of antiparallel spin alignments^{10–12}. Its application to studies of the f.c.c. Fe–Ni system suggests that multiple magnetic states should be of importance for this alloy. Similar conclusions have been drawn from recent experimental work^{15,16}. Also, we note that the Fe–Ni alloy shares some features with γ -Fe. Studies of the latter reveal many different non-collinear states with very similar energies, many of which are more stable than any collinear configuration^{17–20}. A similar situation has been reported for the ordered Fe_3Pt Invar compound¹⁸. But in general, with a few exceptions²¹, *ab initio* calculations of the ground-state properties of random Invar alloys have neglected the possibility of a non-collinear spin alignment. Therefore the continuous nature of the magnetic transition has not previously been discussed.

Figure 1 shows the calculated ground-state spin configurations at selected atomic volumes Ω . The transition from a high-spin ferromagnetic configuration at large volumes to an increasingly disordered non-collinear state is clearly seen as the volume decreases. In Fig. 2 these results are analysed by means of spin pair correlation functions for Fe–Fe and Ni–Ni nearest neighbours. There is a transition from a nearly ferromagnetic ordering (δ -function at $\cos(\theta) = 1$) in Fe–Fe nearest-neighbour (NN) spins at large volumes, to an approximately uniform distribution of spins for small volumes. In contrast, the NN Ni–Ni spins remain nearly

ferromagnetically aligned along the magnetization direction at all volumes. The second NN Fe–Fe and NN Fe–Ni spins are not shown but are similar to the NN Ni–Ni, though with a somewhat larger dispersion than the Ni–Ni case. At $\Omega = 75.7$ a.u. (a.u. is a_0^3 , where a_0 is the Bohr radius), two Fe spins make a discontinuous transition from the ferromagnetic configuration to an approximately anti-ferromagnetic alignment with a slightly reduced local moment. These spin flips catalyse the transition to a non-collinear alignment for smaller lattice constants, and excepting this case, all the remaining spins evolve continuously both in amplitudes and orientations. The spin pair correlation functions (Fig. 2) show that the spins arrange themselves to relax the frustration among the Fe neighbours, while preserving as much as possible a normal ferromagnetic ordering among all other pairs.

As a result of such a continuous transformation of the ground-state spin structure, the average magnetization, $\Sigma_i \mathbf{M}_i \cdot \boldsymbol{\mu}_i$ (shown in Fig. 3a by hexagons), changes much faster than the corresponding averages of the absolute values of local moments of Fe and Ni (Fig. 3a, open circles). (Here, $\mathbf{M}$ is the magnetization direction, and $\boldsymbol{\mu}_i$ is the magnetic moment of atom i .) Thus, the extinction of the magnetization results both from changes in the magnetic angles and the local moments. The behaviour of the average magnetic moment observed in our fully relaxed supercell calculations can be con-

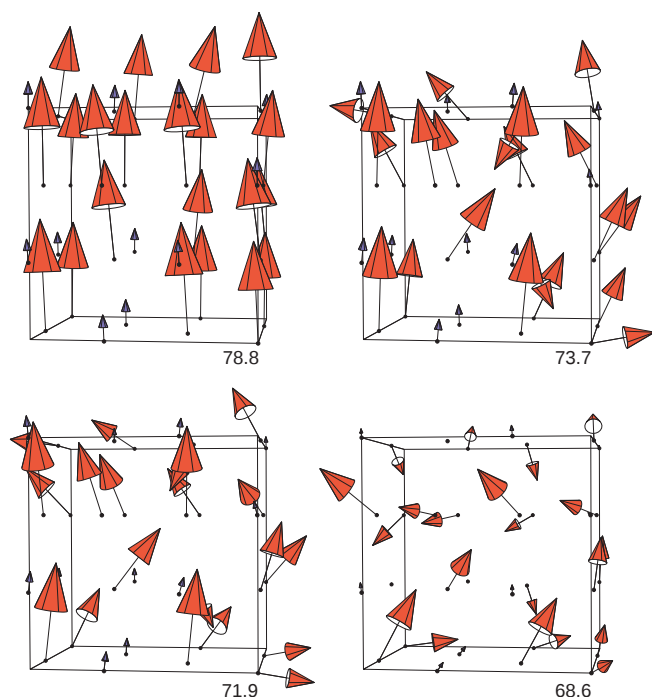


Figure 1 Self-consistent magnetic spin configurations of the f.c.c. Fe–Ni alloy at four different volumes. The volumes (a.u.) are listed on the lower right. Red and blue arrows show magnetic moments on Fe and Ni atoms, respectively. Arrow sizes reflect local moment amplitudes. The Ni spins are always approximately oriented along the average magnetization direction (along the azimuthal direction in the figure). The alloy is modelled by a supercell of 32 atoms with 21 Fe atoms and 11 Ni atoms randomly distributed. The electronic structure was calculated in the framework of the local spin-density approximation²⁶ employing the scalar-relativistic linear muffin-tin orbitals (LMTO) method^{27,28}. Refs 19, 20 describe our treatment of the non-collinear magnetism and the attendant magnetic forces. Calculations were performed with 27 k -points in the Brillouin zone. In the non-collinear case, the density was iterated to self-consistency simultaneously with the relaxation of the magnetic angles. The latter required $\sim 1,000$ band passes for a given lattice constant, at which point the total energy was converged to $\sim 10^{-6}$ Ry per atom. To ensure that the relaxed angles corresponded to a global minimum, we performed ground-state searches starting from several initial angles for a number of the atomic volumes. With one or two exceptions, the same ground state was always found.

trasted with the earlier results of the collinear ferromagnetic calculations for f.c.c. Fe–Ni (refs 2,6–9,13,14); these earlier results show stability of the HS state down to much lower volumes, and a more abrupt transition from the HS to the LS state.

This difference is also reflected in a quite different behaviour of the binding energy curves for the collinear and the non-collinear calculations. In the former case we find (Fig. 3b), in agreement with earlier studies^{6–14,4,5}, that the curve corresponding to the HS ferromagnetic solution (dotted line, open diamonds) crosses the curve for the paramagnetic solution (dotted line, filled diamonds) at a certain volume; this results in the well-known two-branch shape of the binding energy curve. In contrast to the classical view, we find that the transition is continuous (dashed line, filled squares), although the rate of change of magnetization is rapid near the equilibrium volume, reminiscent of the dichotomy in the HS–LS states found by the ferromagnetic collinear calculations. In addition, we show in Fig. 3b (filled circles) results of collinear calculations that allow for both parallel and antiparallel spin alignments. We see that for $\Omega < \sim 74$ a.u. this solution becomes energetically more stable than the ferromagnetic solution. This is in agreement with the conclusions of refs 10–12, where the disordered local moment solution is also found to be more stable than the ferromagnetic solution. But Fig. 3b clearly demonstrates that at low volumes, the non-collinear ground state is indeed more stable than any collinear spin state.

We now discuss the consequences of our results. First, the existence of a stable non-collinear ground state implies that the magnetic moment as a function of concentration will not drop abruptly from a high value to zero, but will decrease gradually, as experimentally observed. Second, the equilibrium volume for the

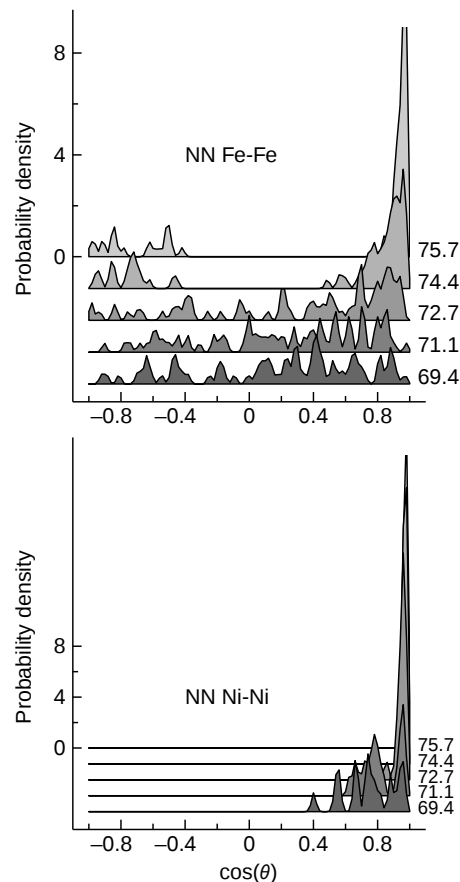


Figure 2 Spin pair correlation functions between neighbouring spins in f.c.c. Fe–Ni at several atomic volumes (in a.u.). Correlation functions $\hat{\mathbf{S}}_i \cdot \hat{\mathbf{S}}_j$ are shown for nearest neighbour (NN) Fe–Fe (panel a) and Ni–Ni (panel b). For larger volumes, $\hat{\mathbf{S}}_i \cdot \hat{\mathbf{S}}_j \approx \delta(\cos\theta - 1)$. Here, $\mathbf{S}_i$ is the direction vector for the spin at site i .

non-collinear alloy is situated between the ferromagnetic HS and paramagnetic states. In particular, the difference between the HS and the non-collinear equilibrium volumes is $\sim 1\%$, which is in excellent agreement with experiment (0.5–2%; refs 13, 22), in contrast to the volume difference between the paramagnetic and the HS ferromagnetic states ($\sim 7\%$). Third, the stability of the non-zero transverse component in the magnetic ground state of the Invar alloy ($T = 0$ K) probably explains the anomalies in the temperature-dependence of the magnetic moments in these systems, instead of an explanation offered by Ishikawa *et al.*²³, who suggested an existence of some “hidden excitation” in Invar systems. Although a quantitative analysis would require spin-dynamical simulations^{19,20} beyond the scope of this Letter, a qualitative explanation might be as

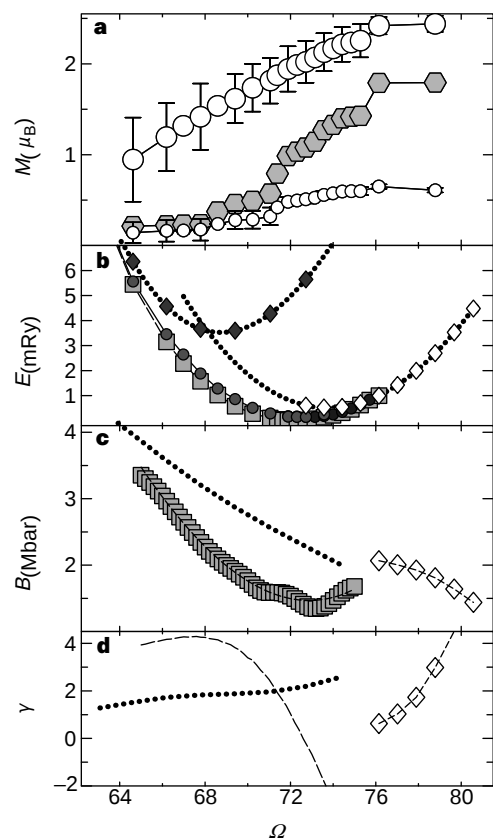


Figure 3 Volume-dependence of ground-state properties of the $\text{Ni}_{0.34}\text{Fe}_{0.66}$ alloy. Atomic volume is in a.u. **a**, Filled hexagons denote average magnetization per atom (in μ_B); large and small open circles represent the average absolute magnetic moments on the Fe and Ni sites, respectively. Error bars denote r.m.s. fluctuations in the moments. **b**, Binding energy curve (in mRy). Dashed line and filled squares, ground-state energy from the non-collinear calculations. Dotted line, and open and filled diamonds, ferromagnetic and non-magnetic calculations, respectively, showing the classical high-spin to low-spin transition. Solid line and filled circles, collinear calculation. **c**, Bulk modulus (Mbar) obtained from differentiation of $E(\Omega)$. Open diamonds and filled squares show the (ground-state) ferromagnetic and non-collinear data, respectively. The ‘wiggles’ for $\Omega < 72$ a.u. are an artefact of slightly inexact forces, which produce a small inconsistency ($\sim 5 \times 10^{-6}$ Ry per atom) between the calculated and true ground-state energy. The dashed line shows a smoothed version of the spline fit using an arbitrary form that approximately preserves the near-linear behaviour and the area under the curve, which must be conserved if the low- and high-volume limits of the pressure are correct. Dotted line, B for the paramagnetic case. **d**, Grüneisen constant γ_{LT} as calculated from equation (1), by differentiating B . The line types are similar to that in **c**. Owing to the ‘wiggles’ in B , shown in **c**, it is not possible to extract a precise value for the Grüneisen parameter that is independent of the fitting procedure. However, the general features (of $\gamma > \sim 3$ for smaller-than-equilibrium volumes, $\gamma < \sim 0$ near the equilibrium volume, and γ rising again for still larger volumes) are definitive.

follows. Starting from the frustrated non-collinear ground state, there is a larger phase space accessible to the system to increase its entropy for a fixed cost in energy, and therefore at low temperatures one should expect substantial deviation in the magnetization $M(T)$ from the Brillouin behaviour found in conventional ferromagnets.

Finally, we discuss the thermal expansion coefficient, the key property of the Invar effect. In general, it is connected to the anharmonicity of the phonons, most importantly the binding energy curve. In particular, this implies that the bulk modulus for a conventional paramagnetic system $B = V \partial^2 E / \partial V^2$ increases with decreasing volume, as shown in Fig. 3c for the case of the paramagnetic alloy (dotted line). (Here, V is the volume and E is the internal energy.) Moruzzi *et al.*²⁴ encapsulated the predominant effects by the so-called low-temperature Grüneisen constant γ_{LT} obtained from the volume dependence of the pressure:

$$\gamma_{\text{LT}} = -1 - \frac{V}{2} \frac{P}{\partial P / \partial V} \quad (1)$$

A typical value of γ_{LT} for paramagnetic transition metals ranges from 1.5 to 2.5 (see dotted line in Fig. 3d for the paramagnetic case) which corresponds to thermal expansion coefficients of the order of 10^{-5} K^{-1} . The magnetic phase transition leads to a substantial softening of the lattice, as is clearly seen from the bulk modulus in Fig. 3c. Indeed, in agreement with experiment, the calculated bulk modulus B with fully relaxed magnetic moments is 1.5 Mbar at the equilibrium volume. We note that this is a much smaller value than one would expect simply extrapolating the paramagnetic or ferromagnetic curves in Fig. 3c towards the equilibrium volume. Thus, the softening originates in the system’s ability to relax its magnetic structure. Consequently, the bulk modulus first decreases from its value in a ferromagnetic system, then increases towards its value in a paramagnetic system, leading to a volume dependence with a well defined minimum. This result is in agreement with the experimental study of Mañosa *et al.*²⁵ who found that the hydrostatic-pressure derivatives of the elastic-stiffness-tensor components in $\text{Fe}_{65}\text{Ni}_{35}$ are positive but very small, and in disordered $\text{Fe}_{72}\text{Pt}_{28}$ Invar they are even negative. The anomalous volume-dependence of B makes it immediately evident that γ_{LT} should also be anomalous. Indeed, when γ_{LT} is extracted according to equation (1), it is anomalously large for small volumes, and is ~ 0 for the equilibrium volume. Thus the anomalous parameters of the binding energy curve emerges as a straightforward result of the magneto-volume effect (that is, coupling of the magnetic to the elastic energies). □

Received 3 February; accepted 4 May 1999.

- Guillaume, C. E. Recherches sur les aciers au nickel. Dilatations aux températures élevées; résistance électrique. *CR Acad. Sci.* **125**, 235–238 (1897).
- Wasserman, E. F. in *Ferromagnetic Materials* Vol. 5 (eds Buschow, K. H. J. & Wohlfarth, E. P.) 237–322 (North-Holland, Amsterdam, 1990).
- Weiss, R. J. The origin of the ‘Invar’ effect. *Proc. R. Soc. Lond. A* **82**, 281–288 (1963).
- Andersen, O. K., Madsen, J., Poulsen, U. K., Jepsen, O. & Kollar, J. Magnetic ground state properties of transition metals. *Physica B* **86–88**, 249–256 (1977).
- Roy, D. M. & Pettifor, D. G. Stoner theory support for the two-state hypothesis for γ iron. *J. Phys. F* **7**, L183–L187 (1977).
- Moroni, E. G. & Jarlborg, T. Calculation of Invar anomalies. *Phys. Rev. B* **41**, 9600–9602 (1990).
- Moruzzi, V. L. High-spin and low-spin states in Invar and related alloys. *Phys. Rev. B* **41**, 6939–6946 (1990).
- Mohn, P., Schwarz, K. & Wagner, D. Magnetoelastic anomalies in Fe–Ni Invar alloys. *Phys. Rev. B* **43**, 3318–3324 (1991).
- Entel, P., Hoffmann, E., Mohn, P., Schwarz, K. & Moruzzi, V. L. First-principles calculations of the instability leading to the Invar effect. *Phys. Rev. B* **47**, 8706–8720 (1993).
- Johnson, D. D., Pinski, F. J., Staunton, J. B., Györfy, B. L. & Stocks, G. M. in *Physical Metallurgy of Controlled Expansion Invar-Type Alloys* (eds Russell, K. C. & Smith, D. E.) 3–24 (The Minerals, Metals & Materials Soc., Warrendale, Pennsylvania, 1990).
- Akai, H. & Dederichs, P. H. Local moment disorder in ferromagnetic alloys. *Phys. Rev. B* **47**, 8739–8747 (1993).
- Johnson, D. D. & Shelton, W. A. in *The Invar effect: A Centennial Symposium* (ed. Wittenauer, J.) 63–74 (The Minerals, Metals & Materials Soc., Warrendale, Pennsylvania, 1997).
- Abrikosov, I. A., Eriksson, O., Söderlind, P., Skriver, H. L. & Johansson, B. Theoretical aspects of the Fe–Ni–Invar alloys. *Phys. Rev. B* **51**, 1058–1063 (1995).
- Schröter, M. *et al.* First-principles investigations of atomic disorder effects on magnetic and structural instabilities in transition-metal alloys. *Phys. Rev. B* **52**, 188–209 (1995).
- Schumann, F. O., Willis, R. F., Goodman, K. G. & Tobin, J. G. Magnetic instability of ultrathin fcc $\text{Fe}_2\text{Ni}_{1-x}$ films. *Phys. Rev. Lett.* **79**, 5166–5169 (1997).
- Freeland, J. W., Grigorov, I. L. & Walker, J. C. Magnetic phase transition in epitaxial $\text{Ni}_{1-x}\text{Fe}_x$ alloy thin films. *Phys. Rev. B* **57**, 80–83 (1998).

17. Mryasov, O. N., Gubanov, A. V. & Liechtenstein, A. I. Spiral spin-density-wave states in FCC iron: Linear-muffin-tin-orbitals band-structure approach. *Phys. Rev. B* **45**, 12330–12336 (1992).
18. Uhl, M., Sandratskii, L. M. & Kübler, J. Spin fluctuations in γ -Fe and in Fe₃Pt Invar from local-density-functional calculations. *Phys. Rev. B* **50**, 291–301 (1994).
19. Antropov, V. P., Katsnelson, M. I., van Schilfhaarde, M. & Harmon, B. N. Ab initio spin dynamics in magnets. *Phys. Rev. Lett.* **75**, 729–732 (1995).
20. Antropov, V. P., Katsnelson, M. I., Harmon, B. N., van Schilfhaarde, M. & Kuznezov, D. Spin dynamics in magnets: equation of motion and finite temperature effects. *Phys. Rev. B* **54**, 1019–1035 (1996).
21. Wang, Y. *et al.* Noncollinear magnetic structure in Ni_{0.35}Fe_{0.65}. *J. Appl. Phys.* **81**, 3873–3875 (1997).
22. Pearson, W. B. *A Handbook of Lattice Spacing and Structure of Metals and Alloys* (Pergamon, London, 1967).
23. Ishikawa, Y., Noda, Y., Ziebeck, K. R. A. & Givord, D. Origin of hidden magnetic excitations in an Invar alloy Fe₆₅Ni₃₅. *Solid State Commun.* **57**, 531–534 (1986).
24. Moruzzi, V. L., Janak, J. F. & Schwarz, K. Calculated thermal properties of metals. *Phys. Rev. B* **37**, 790–799 (1988).
25. Mañosa, L. *et al.* Acoustic-mode vibrational anharmonicity related to the anomalous thermal expansion of Invar iron alloys. *Phys. Rev. B* **45**, 2224–2236 (1992).
26. von Barth, U. & Hedin, L. A local exchange-correlation potential for the spin polarized case: I. *J. Phys. C* **5**, 1629–1642 (1972).
27. Andersen, O. K. Linear methods in band theory. *Phys. Rev. B* **12**, 3060–3083 (1975).
28. Andersen, O. K., Jepsen, O. & Glötzel, D. in *Highlights of Condensed-Matter Theory* (eds Bassani, F., Fumi, T. & Tosi, M. P.) 59–176 (North-Holland, New York, 1985).

Acknowledgements. M.v.S. thanks V. Antropov for discussions; I.A.A. and B.J. were supported by the Swedish Natural Science Council; M.v.S. was supported by ONR. This work was also supported by the Swedish Materials Consortium No. 9.

Correspondence and requests for materials should be addressed to M.v.S. (e-mail: markv@waldo.ran.sandia.gov).

Confinement-induced miscibility in polymer blends

S. Zhu*, Y. Liu*, M. H. Rafailovich*, J. Sokolov*, D. Gersappe*, D. A. Winesett† & H. Ade†

* Department of Materials Science and Engineering, State University of New York, Stony Brook, New York 11794, USA

† Department of Physics, North Carolina State University, Raleigh, North Carolina 27695, USA

The use of polymer thin films in technology is increasingly widespread—for example, as protective or lithographic surface coatings, or as active (electronic or optical) elements in device architectures. But it is difficult to generate films of polymer mixtures with homogeneous surface properties, because of the tendency of the polymers to phase-separate^{1,2}. Copolymer compatibilizers can induce miscibility in polymer blends, but only with chemical components that are either close to a critical point in the phase diagram³ or which have an attractive interaction between them^{4,5}. Instead of manipulating the chemical composition of the blend, we show here that complete mixing can be obtained in polymer blends by the physical effect of confinement in thin films. The compatibilization results from entropic inhibition of phase separation into micelles, owing to confinement. The result is an intimately mixed microemulsion with a perfectly flat surface and a two-dimensional maze-like structure with columnar domains that extend through the film.

The critical factor that controls the ability of a copolymer compatibilizer to improve the miscibility in polymer blends is the balance between the chemical potentials of the copolymer localized at the interface and the copolymer in a micelle formed in one of the homopolymer phases. Earlier studies have shown that in the bulk, this balance always shifts to micelle formation before the interface can get saturated with copolymer, thus limiting the miscibility of the blend^{6–10}. To apply the principles of compatibilization to thin film blends, we need to account for the effects that the confinement due to the thin film geometry will have on this chemical potential balance. We approach this problem in the following stages. First, we use a theoretical mean-field model to identify the factors that control the process of micelle formation in confined systems and set guidelines for inducing complete miscibility in thin film polymer blends. Then, we proceed to test the theoretical predictions by performing a series of experiments on thin film polymer blends.

We use a numerical self-consistent field (SCF) model based on the method developed by Scheutjens and Fleer¹¹. The SCF method studies polymeric phase behaviour by combining markovian statistics with a mean-field approximation. In this model, the free energy of the system is:

$$F = k\ln\Omega + \sum \chi\phi_i(r)(1 - \phi_i(r)) \quad (1)$$

where the first term represents the entropy of the system, and the second term represents the enthalpy of the system. Here k is Boltzmann's constant, Ω is the number of conformations possible, χ is the Flory–Huggins parameter that represents the various monomer–monomer interaction energies and $\phi_i(r)$ is the concentration of species i at position r . The aim of the calculation is to determine a concentration profile that minimizes the free energy of the system.

We focus on linear symmetric diblock copolymers, as these copolymers have been shown to pack more efficiently at the interface than comb, star or random copolymers¹². Therefore, if we understand how to retard micelle formation in diblock copolymers, these copolymers will be the most efficient at reducing the interfacial tension between the polymer phases. In our model we assume that the micelles formed are spherical, because studies have shown that due to kinetic effects, spherical micelles are the primary micelle structure over a wide range of copolymer concentrations¹³.

In Fig. 1a we plot the critical micelle concentration (c.m.c.; the concentration at which micelles first appear in the system) for two different diblock copolymers as a function of the degree of

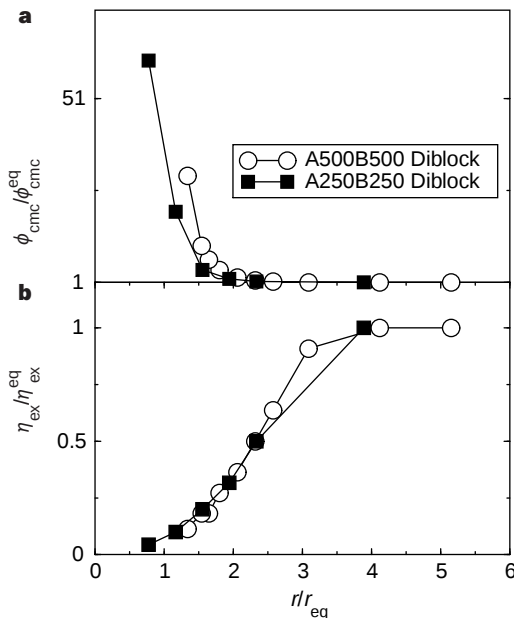


Figure 1 Results of the SCF calculations showing the effect of confinement on micelle formation. **a**, Plot of the c.m.c. of the confined system, ϕ_{cmc} , scaled by the bulk c.m.c., ϕ_{cmc}^{eq} , as a function of the degree of confinement. **b**, Plot of the size of the confined micelle, η_{ex} , scaled by the equilibrium bulk micelle size η_{ex}^{eq} . In **a** and **b**, r/r_{eq} refers to the radius of the equilibrium bulk micelle, and r is the system size. We considered the formation of spherical micelles by linear symmetric diblock copolymers in a homopolymer phase by using a lattice with spherical symmetry. To model the confinement we bounded the system by a hard, neutral wall a distance r away from the centre of the lattice (by decreasing r we increase the confinement of the micelle). We considered two diblocks with different lengths, a A250B250 (250 segments of A monomers followed by 250 segments of B monomers) diblock and a A500B500 diblock. In all cases, the homopolymer phase was a B homopolymer of 100 lattice units in length. The energetic parameter χ_{AB} was set at 0.1.

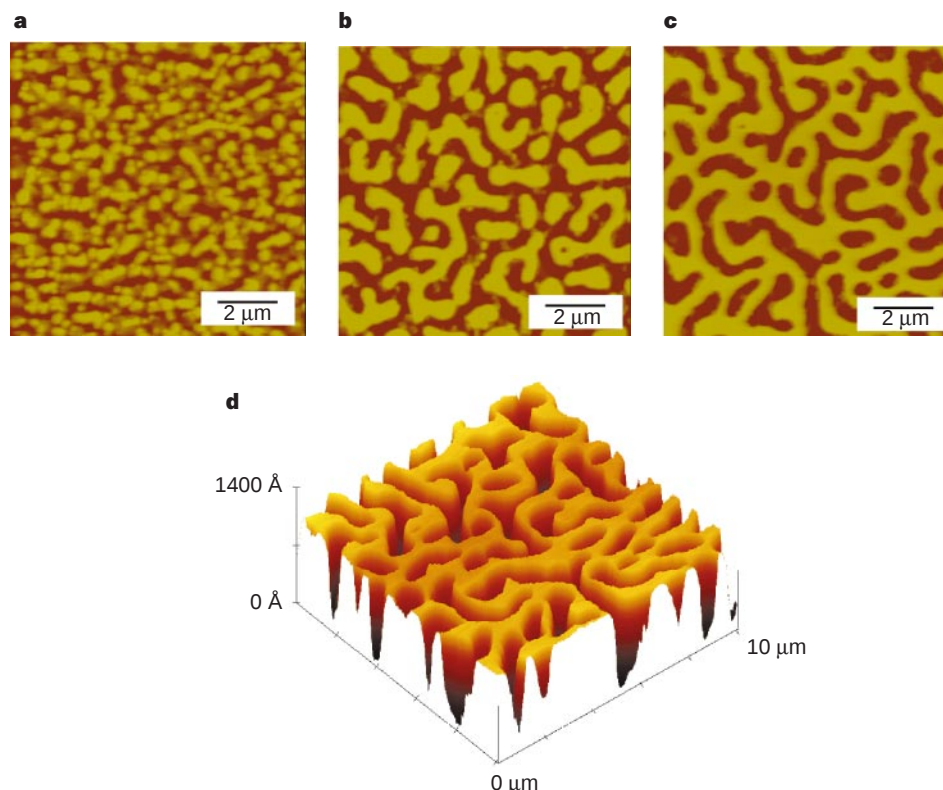


Figure 2 SFM topographical images of bilayer samples. The samples were annealed for seven days at 180 °C in 10^{-7} torr vacuum and washed in cyclohexane to selectively dissolve the PS homopolymer. **a**, When the thickness of the PS layer, t , is 1,000 Å, or t is greater than the micelle diameter, we observe undissolved micelles that remain on the PMMA layer surface. **b**, Intermediate

region, $t = 800$ Å, where coexisting micelles and microemulsion structures are observed. **c**, Confined regime where $t = 500$ Å or t is less than the micelle diameter, and only a microemulsion is seen. **d**, Three-dimensional view of **c**, showing that the structures are columnar and extend the whole height of the original bilayer.

confinement. We see that large changes in the c.m.c. only happen at high degrees of confinement. Initially, at small degrees of confinement, the micelle reacts by undergoing a slight decrease in its size (Fig. 1b). This change in size, however, is not large enough to affect the free energy of micelle formation. As we increase the confinement further, the micelle gets smaller and the entropic losses that occur with the packing of linear diblock copolymers into smaller and smaller spheres starts increasing rapidly. Consequently, the formation of micelles becomes energetically unfavourable, and this shows up in a large increase in the c.m.c. of the system.

The length scale at which confinement changes the behaviour of the system is approximately twice the radius of the bulk micelle (Fig. 1a). It is important to note, however, that even though the length scale appears related to the bulk micelle size, at this point the size of the micelle is roughly half the size of the original bulk micelle. After this point, the c.m.c. of the system increases very rapidly and consequently so does the chemical potential of the copolymers in a micelle. Thus, confinement makes the copolymer localization at the interface more favourable than the formation of micelles, and there is now a strong driving force for the chains to migrate to the interface between the homopolymers. At high enough concentrations of copolymer, this migration should result in a vanishing interfacial tension and in the formation of a microemulsion phase.

To test the theoretical prediction we chose an experimental system consisting of two well characterized monodisperse (that is, the polydispersity index is < 1.1) polymers, polystyrene (PS) and polymethylmethacrylate (PMMA). The molecular masses of PS and PMMA were 200,000 and 296,000 daltons, respectively. We chose this system because these homopolymer pairs are highly incompatible (the χ parameter between the homopolymers is ~ 0.04) and in

this strongly segregated system, the formation of a microemulsion has never been observed^{9,10}. The diblock used in this study was a 37,000–46,000-dalton deuterated PS (dPS)-PMMA copolymer. Previous studies have established that in bulk systems, this copolymer forms spherical micelles in the PS phase¹⁴. Using an analytical expression derived by Semenov¹³, we estimated the diameter of these micelles to be 560 Å. All the samples were prepared in the following manner. First, on a cleaned Si substrate we spun-cast a thick layer of PMMA (> 570 Å). On it, we floated a layer that contained PS and 30% copolymer dPS-PMMA—a concentration which is much larger than the c.m.c. of the diblock¹⁴. We control the degree of confinement in our system by controlling the thickness of the top PS layer. To reach equilibrium, all the samples were annealed in a vacuum of 10^{-7} torr at 180 °C for times ranging from 24 hours to a week.

In Fig. 2, we show scanning force microscopy (SFM) images of the bilayer samples. For thick PS layers (> 900 Å), we can see the micelles that remain after the dissolution of the PS phase on the surface of the PMMA layer (Fig. 2a). When the PS layer thickness was then gradually decreased to less than the size of the bulk micelle diameter (< 560 Å), the point where the theory predicts a large change in the c.m.c. of the system, the two homopolymers completely mixed (Fig. 2c). From Fig. 2d we can see that the resulting structure is a microemulsion with columnar domains that extend through the entire film. This maze-like configuration has a characteristic period which corresponds to a domain spacing of 600 nm.

To make sure that we were not influencing the morphology of the system by our washing procedure, we also performed near edge X-ray adsorption fine structure (NEXAFS) microscopy scans of the thin-film blend samples. NEXAFS microscopy is a technique by which we can obtain two-dimensional composition profiles of thin

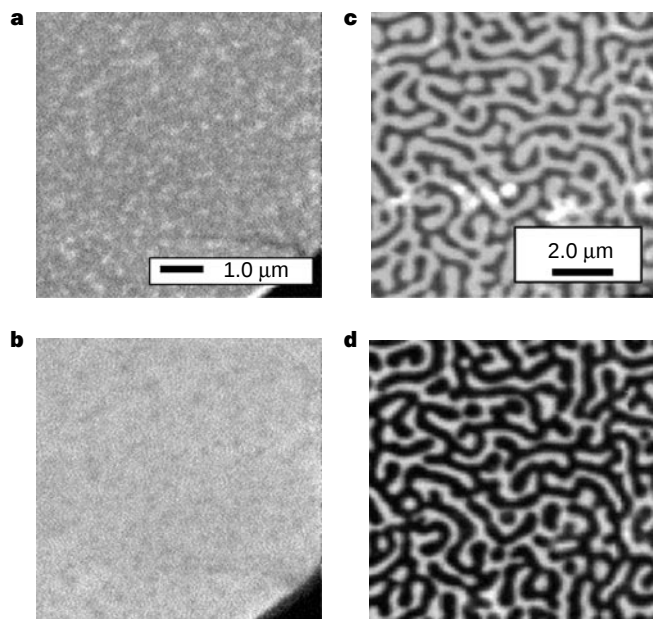


Figure 3 NEXAFS microscopy images showing the effect of the PS film thickness on morphology after annealing. **a, c** and **b, d** are the PS and PMMA distributions, respectively. **a, b**, 1,100-Å PS floated onto 570-Å PMMA and annealed for 24 h, showing the presence of micelles; **c, d**, 380-Å PS floated on 950-Å PMMA and annealed for 96 h, showing the microemulsion phase. Note that the domain spacings after annealing for 96 h are similar to that observed after 7 d of annealing, indicating that the structures are stable in time. Here images **a** and **b** are on the same scale, and images **c** and **d** are on the same scale.

films with a spatial resolution of 40 nm without any further sample treatment¹⁵. For films where the thickness of the PS layer was 1,100 Å (Fig. 3a and b), spherical micelles can be clearly seen in the system. When the PS layer is decreased to 380 Å (below the diameter of the bulk micelle), a microemulsion phase with a well defined wavelength is observed (Fig. 3c and d). The height of each of the domains in the pattern is approximately equal to the initial thickness of the sample, indicating that the structure extends all the way to the substrate in agreement with the SFM results. In the total density profile, which yields the same information as SFM topography, only a slight roughening of the surface is observed. This is in sharp contrast to observations of bilayers of thin PS on PMMA¹⁴ (without compatibilizer), where the PS formed droplets (with well defined contact angles) on the PMMA, rather than wetting it. Thus we believe that what we are observing here is the formation of a true microemulsion phase, and not a phase-separated system. We note that the NEXAFS microscopy will allow us to monitor the dynamics of microemulsion formation, and work on this is in progress.

To confirm the location of the diblock copolymer at the interface between the PS-PMMA domains, we washed the sample that formed the microemulsion with cyclohexane and scanned it using the lateral force mode of the SFM. The regions with higher friction contrast are located at the edges of the PMMA domains (Fig. 4). These are due to the PS block of the diblock copolymer which, unlike the PS homopolymer, does not readily dissolve in cyclohexane.

As our approach does not restrict the chemical nature of the polymers, it should be able to produce complete mixing in any system in which the formation of micelles limits the miscibility of the blend. We therefore expect that this method will have a significant effect on any technological process that relies on ultra-thin polymer coatings, such as photolithographic printing and

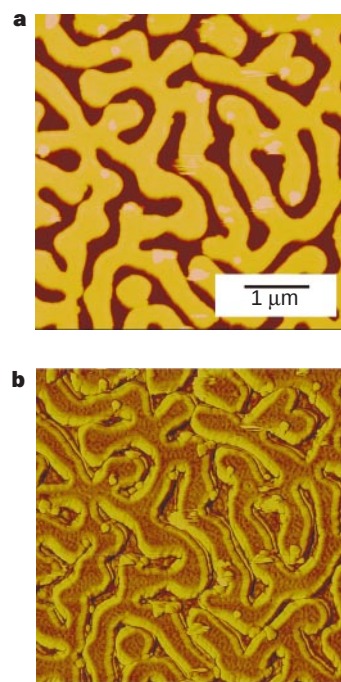


Figure 4 SFM images showing the location of diblocks at the interface between the two homopolymers. **a**, SFM topography, and **b**, friction force microscopy, on sample shown in Fig. 3c and d after dissolution of the PS component with cyclohexane. Note the friction contrast at the interface of the PMMA domains, indicating the presence of the diblock copolymer. Here images **a** and **b** are on the same scale.

magnetic-disk coatings, as our technique produces films that are perfectly flat and completely mixed. □

Received 12 February; accepted 20 April 1999.

- Karim, A. *et al.* Phase-separation induced surface patterns in thin polymer blend films. *Macromolecules* **31**, 857–862 (1998).
- Dalnoki-Veress, K., Forrest, J. A., Stevens, J. R. & Dutcher, J. R. Phase separation morphology of spin-coated polymer blend thin films. *Physica A* **239**, 87–94 (1997).
- Bates, F. S. *et al.* Polymeric bicontinuous microemulsions. *Phys. Rev. Lett.* **79**, 849–852 (1997).
- Shull, K. R., Kellock, A. J., Deline, V. R. & Macdonald, S. A. Vanishing interfacial-tension in an immiscible polymer blend. *J. Chem. Phys.* **97**, 2095–2104 (1992).
- Xu, Z., Jandt, K. D., Kramer, E. J., Edgecombe, B. D. & Frechet, J. M. J. Direct observation of a diblock copolymer-induced microemulsion at a polymer/polymer interface. *J. Polym. Sci. B* **33**, 2351–2357 (1995).
- Shull, K. R. Interfacial phase-transitions in block copolymer homopolymer blends. *Macromolecules* **26**, 2346–2360 (1993).
- Shull, K. R., Kramer, E. J., Hadziannou, G. & Tang, W. Segregation of block copolymers to interfaces between immiscible homopolymers. *Macromolecules* **23**, 4780–4787 (1990).
- Hu, W. C., Koberstein, J. T., Lingel, J. P. & Gallot, Y. Interfacial-tension reduction in polystyrene/poly(dimethylsiloxane) blends by the addition of poly(styrene-*b*-dimethylsiloxane). *Macromolecules* **28**, 5209–5214 (1995).
- Russell, T. P. *et al.* Width of homopolymer interfaces in the presence of symmetrical diblock copolymers. *Macromolecules* **24**, 5721–5726 (1991).
- Green, P. F. & Russell, T. P. Segregation of low-molecular-weight symmetrical diblock copolymers at the interface of high-molecular-weight homopolymers. *Macromolecules* **24**, 2931–2935 (1991).
- Fleer, G., Cohen-Stuart, M. A., Scheutjens, J. M. H. M., Cosgrove, T. & Vincent, B. *Polymers at Interfaces* (Chapman & Hall, London, 1993).
- Lyatskaya, Y., Gersappe, D., Gross, N. A. & Balazs, A. C. Designing compatibilizers to reduce interfacial tension in polymer blends. *J. Phys. Chem.* **100**, 1449–1458 (1996).
- Semenov, A. N. Theory of diblock-copolymer segregation to the interface and free-surface of a homopolymer layer. *Macromolecules* **25**, 4967–4977 (1992).
- Qu, S. *Dewetting Dynamics at Polymer Interfaces*. Thesis, State University of New York at Stony Brook (1998).
- Ade, H. *et al.* Bulk and surface characterization of a dewetting thin film polymer bilayer. *Appl. Phys. Lett.* **73**, 3775–3777 (1998).

Acknowledgements. NEXAFS microscopy was performed with the Stony Brook STXM at the National Synchrotron Light Source, developed by the groups of J. Kirz and C. Jacobsen with support from the Office of Biological and Environmental Research, and the NSF. We thank S. Spector and C. Jacobsen for developing the zone plates, with support from the NSF. The NSLS is supported by the Office of Basic Energy Sciences, Energy Research, Department of Energy. This work was supported by the NSF DMR-MRSEC Program and by a NSF Young Investigator award.

Correspondence and requests for materials should be addressed to D.G. (e-mail: dilip.gersappe@sunysb.edu).

Spontaneous assembly of ten components into two interlocked, identical coordination cages

Makoto Fujita^{*†}, Norifumi Fujita[‡], Katsuyuki Ogura[§] & Kentaro Yamaguchi^{||}

^{*} Coordination Chemistry Laboratories, Institute for Molecular Science, and CREST, Japan Science and Technology Corporation (JST), Myodaiji, Okazaki 444-8585, Japan

[‡] Graduate University for Advanced Studies, Myodaiji, Okazaki 444-8585, Japan

[§] Department of Applied Chemistry, Faculty of Engineering, Chiba University, Yayoicho, Inageku, Chiba 263-8522, Japan

^{||} Chemical Analysis Center, Chiba University, Yayoicho, Inageku, Chiba 263-8522, Japan

Supermolecules consisting of interlinked ring-like molecules (catenanes)¹ are an interesting target for chemical synthesis both for their intrinsic interest as non-covalently bound but robust assemblies and because of the perspective they offer on materials chemistry². Catenanes have been prepared by metal-ion templating^{3,4} and self-assembly through other non-covalent interactions^{5–12}. Here we report the synthesis of a catenane composed not of two interlocking rings but of two cages. This structure is prepared by metal-mediated self-assembly^{13–15}. The framework of each cage is assembled from five components: two tridentate ligands held together with three metal ions. Because each cage framework can bind an aromatic ring, two cage units will bind one another during their assembly process through the formation of a quadruple aromatic stack, giving rise to the ten-component interlocked supermolecule^{16,17}.

Our strategy for constructing the interlocked-cage complex is revealed by the retrosynthetic analysis of a target framework (Fig. 1b). First, the interlocked cages (A, in Fig. 1a) are divided into two separate cages (B), in which labile metal–ligand bonds are put to make a reversible pathway ($A \rightleftharpoons B$). Second, the metal-containing cage (B) is deformed so that it has a flat ‘floor’ and a ‘ceiling’ (C). The distance between the ‘floor’ and the ‘ceiling’ should be adjusted so that another copy of B would fit in the cavity. Thus, cage-like complex 1 is designed, in which the interplanar surface-to-surface distance between the ‘floor’ and the ‘ceiling’ is ~ 3.5 Å, an ideal distance for binding an aromatic ring. Finally, this cage compound 1 is disconnected into three component molecules 2–4. *Cis*-protected Pd(II) and Pt(II) complexes 2 have been employed for metal-mediated self-assembly of discrete structures^{14,15}. Quite recently, rectangular molecular boxes with interplanar separation of ~ 3.5 Å were shown to effectively self-assemble into catenated dimers¹⁸.

In line with this strategy, we combined the three components 2a, 3 and 4 in a 3:1:1 stoichiometry in D₂O. At first, a kinetically distributed mixture was formed, which, however, gradually turned into a thermodynamically stable structure after heating the solution at 100 °C for 3 d. This slow conversion is due to the dual character of the Pt(II)–pyridine coordinate bond which is inert at room temperature but which becomes labile at elevated temperatures¹⁹. The thermodynamically favourable structure therefore self-assembles only at elevated temperatures. Nuclear magnetic resonance (NMR) of the crude reaction solution showed the formation of almost a single component, which was isolated as a pure form (65%). After being converted to a salt of PF₆[−], the product was

assigned as 5a (PF₆[−] salt) by electrospray ionization mass spectrometry (ESI-MS). Similarly, the Pd(II) counterpart 5b was prepared, whose formula was confirmed by elemental analysis.

Reliable evidence for the interlocked-cage structure of 5a is obtained by X-ray crystallographic analysis (Fig. 2), which shows that 5a consists of two identical cage frameworks interlocking with each other. The crystal system was assigned to be C-centred monoclinic (C2/c) and the crystallographically solved structure contained only one cage framework, which gave the interlocked structure after symmetry operations. From the side view of 5a (Fig. 2b), the efficient quadruple stacking of aromatic rings is observed. This stacking makes the interlocked structure the most stable of any possible structures formed by the complexation of 2 with 3 and/or 4. From the top view (Fig. 2b), it is found that the

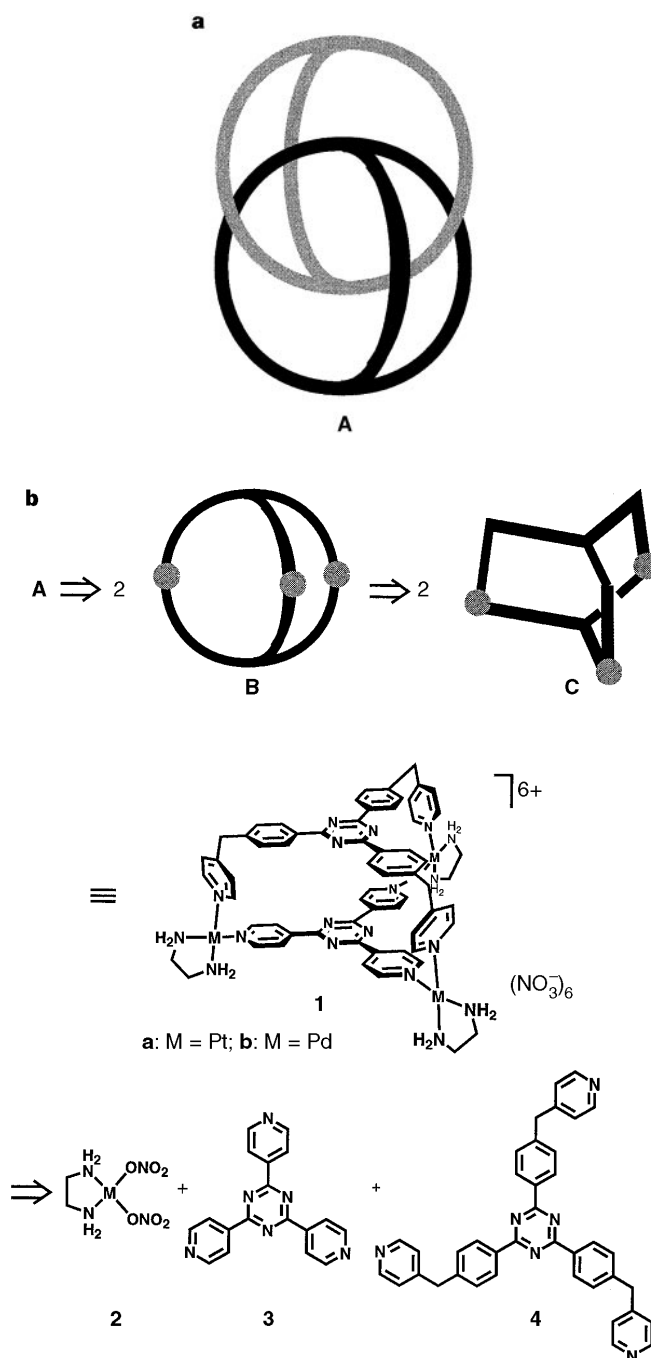


Figure 1 Assembly of the interlocked-cage complex. **a**, Schematic representation of the interlocked supermolecule **A**. **b**, Retrosynthetic analysis.

[†] Present address: School of Engineering, Nagoya University, Chikusa-ku, Nagoya 464-8603, Japan.

triazine core rings of the two cage units are slightly displaced by ~ 1.3 Å, probably due to efficient π - π stacking^{20,21} or favourable crystal packing. We note that interlocked cages are further stacked in an intermolecular fashion, leading to one-dimensional infinite stacking of triazine π -systems. Experimentally, a single crystal of the BF_4^- salt of **5a** was obtained by adding an excess amount of NaBF_4 to the aqueous solution of **5a**, and allowing the solution to stand at ambient temperature for a few weeks.

The structure of **5a,b** in solution was determined by NMR studies. Both ^1H and ^{13}C NMR agree well with the structure of **5a,b** (Fig. 3) and the assignments are fully supported by two-dimensional NMR measurements (see Supplementary Information). The interlocked framework is consistent with (1) an outstanding upfield shift of the inside protons of C_6H_4 ($\Delta\delta = (\delta \text{ in } \mathbf{4}) - (\delta \text{ in } \mathbf{5}) = 2.63$) which is the most significantly shielded by another cage framework, and (2) the observation of the nuclear Overhauser

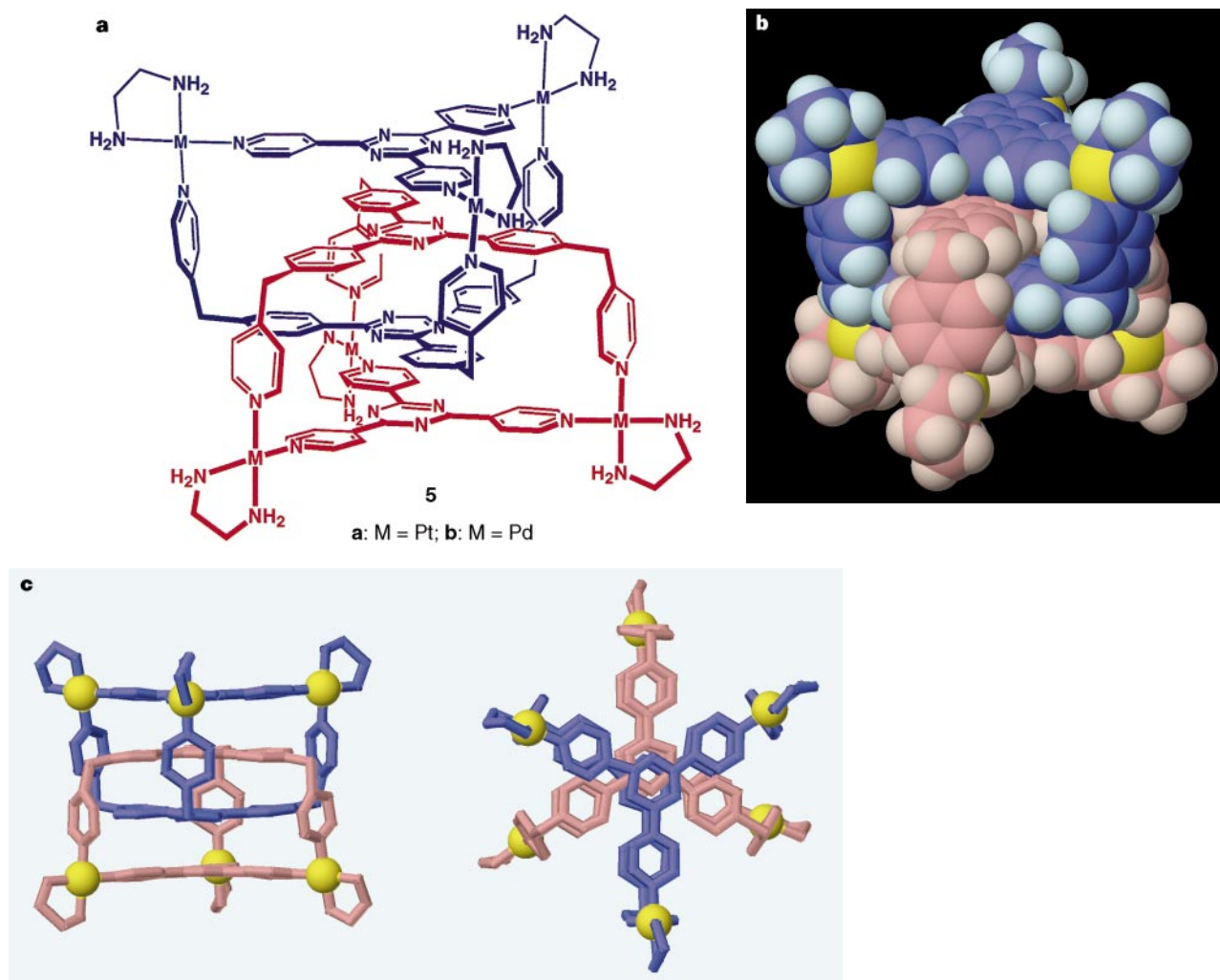


Figure 2 Structure of the interlocked-cage complex. **a**, Structural formula of **5**. **b**, Space-filling presentation of the X-ray crystallographic structure of **5a**. Counter-ions and solvent molecules are omitted for clarity. **c**, A top and a side view of the X-ray structure of **5a** (a ball-and-stick presentation).

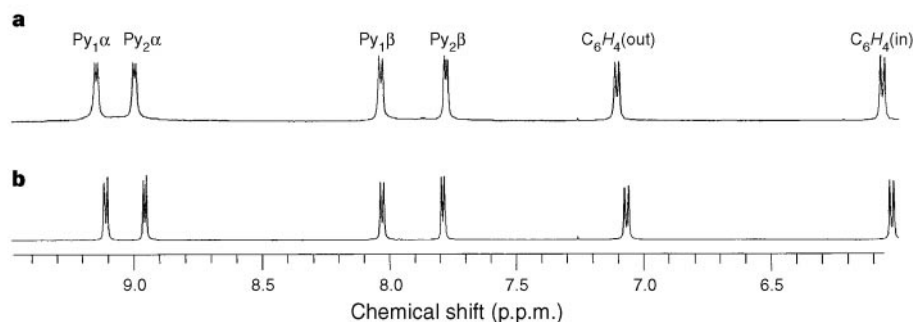


Figure 3 NMR spectra of **5a** (panel **a**) and **5b** (panel **b**) (500 MHz, D_2O , aromatic region). Py₁, pyridine protons of ligand **3**; Py₂, pyridine protons of ligand **4**. NOEs were observed between Py₁α-Py₂α, Py₁β-Py₂α, ArCH₂-Py₂α, C₆H₄(in)-Py₂α, C₆H₄(out)-Py₂α, Py₁α-Py₂β, Py₁β-Py₂β, ArCH₂-Py₂β, C₆H₄(in)-Py₂β and C₆H₄(out)-Py₂β.

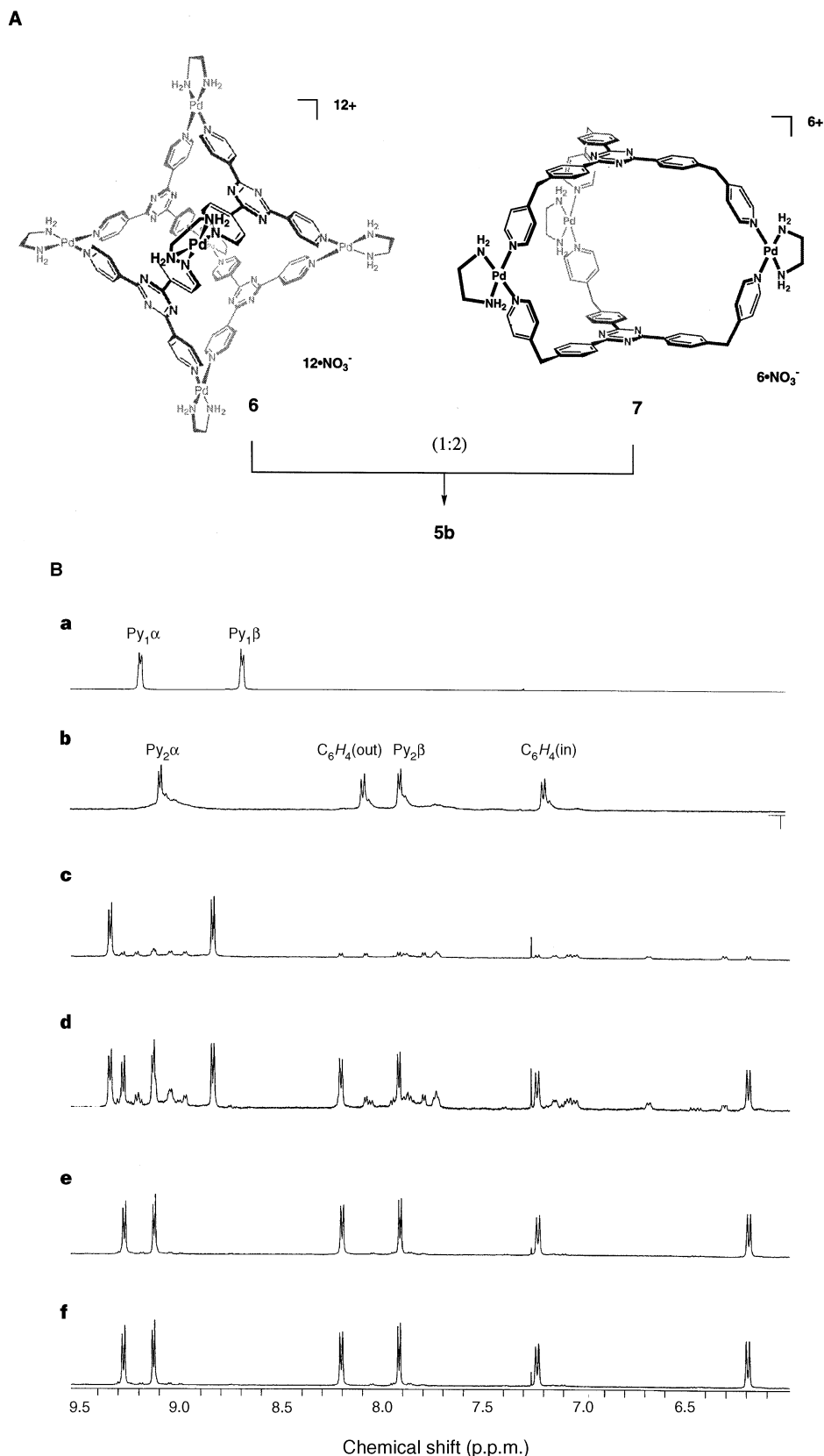


Figure 4 Formation of **5b** via reorganization from **6** and **7**. **A**, The reaction scheme. **B**, NMR observation of the reorganization process. **a**, NMR of **6**: a D_2O solution (1 ml) of **6** was prepared from **3** (2.1 mg, 6.7×10^{-3} mmol) and $\text{enPd}(\text{ONO}_2)$ (0.01 mmol) at 80°C for a day. Py_1 , pyridine protons of ligand **3**. **b**, NMR of **7**: a $\text{D}_2\text{O}-\text{CD}_3\text{OD}$ (1:1) solution (0.5 ml) of **7** was prepared from **4** (3.9 mg, 6.7×10^{-3} mmol) and $\text{enPd}(\text{ONO}_2)$ (0.01 mmol) by heating at 80°C for 5 d. Py_2 , pyridine protons of ligand **4**. The assembly of cage **7** is slow, and oligomeric products still remain

even after 5 d. **c**, The two solutions described above were combined and the NMR spectrum was measured immediately. Cage **6** remained intact, whereas cage **7** turned into oligomeric components due probably to interactions with components of **6**. **d**, Spectrum after 3 h at 20°C . **e**, after 24 h at 20°C . **f**, the two solutions described in **a** and **b** were combined and the NMR spectrum was measured after 10 min at 80°C . Chemical shift values are slightly different from those shown in Fig. 3 because of the measurement in a mixed solvent ($\text{D}_2\text{O}-\text{CD}_3\text{OD}$ (3:1)).

effect (NOE) between $\text{Py}_1\text{-CH}_2\text{-C}_6\text{H}_4$ and Py_2 units (see Fig. 3 legend for definition of Py_1 and Py_2). As the Pt(II)-Py bond is kinetically inert at room temperature, complex **5a** is stable at all concentrations and is not in equilibrium with other structures.

ESI-MS studies also confirmed the structure of **5a** in solution. When an acetonitrile solution of **5a** (PF_6^- salt) was subjected to the ESI-MS, prominent peaks for $[\text{M}-(\text{PF}_6)_n]^{n+} \cdot (\text{CH}_3\text{CN})_m$ ($n = 4-9$, $m = 2-18$) were clearly observed with reasonable isotopic distribution. We note that acetonitrile molecules seem to weakly solvate the cationic sphere of Pt^{2+} ions, and that the number of solvated acetonitrile molecules (m) increases with increasing positive charges (n). Similarly, the ESI-MS study of the BF_4^- salt of **5a** afforded a series of peaks corresponding with $[\text{M}-(\text{BF}_4)_n]^{n+} \cdot (\text{CH}_3\text{CN})_m$.

A striking observation is that two preformed Pd(II) cage compounds **6** and **7** are reorganized into the three-dimensionally interlocked complex **5b** in high yield when they are mixed in an aqueous solution (Fig. 4A). M_6L_4 type complex **6** is a very stable cage compound which assembles from ten components (**6-2b** + **4-3**; ref. 22) (Fig. 4B, a). M_3L_2 type cage **7** is also found to assemble from five components (**3-2b** + **2-4**) concomitant with the formation of oligomeric products (Fig. 4B, b). These cage compounds **6** and **7** were independently prepared in D_2O and then combined in 1:2 ratio so that the component ratio **2b**:**3**:**4** became 6:2:2. The NMR of the mixture initially showed that cage **6** existed intact whereas cage **7** was considerably converted to oligomeric components (Fig. 4B, c). After the solution was allowed to stand at room temperature, cages (**6**, **7**) and oligomers gradually disappeared; alternatively, another set of signals assignable to **5b** appeared (Fig. 4B, d). Finally, all components were completely reorganized into interlocked-cage compound **5b** after 1 d (Fig. 4B, e). At an elevated temperature (80°C), the reorganization (**6** + **2-7** $\rightarrow$ **2-5b**) was significantly accelerated and completed within 10 min (Fig. 4B, f). These results clearly show that **5b** is thermodynamically stable, and is the most stable of all the possible structures formed by combination²³ of components **2b**, **3** and **4**.

We have already reported the guest-induced organization of a three-dimensional cage compound; the host framework was organized from its components only in the presence of a specific guest²⁴. In this context we note that, in the present study, the organization of a cage framework (**3-2** + **3** + **4**) is induced by another copy of itself through interlocking. □

Methods

Self-assembly of 5a. Ligands **3** (28 mg, 0.088 mmol) and **4** (51 mg, 0.088 mmol) were suspended in an aqueous solution (26 ml) of **2a** (100 mg, 0.26 mmol), and the mixture was stirred at 100°C for 3 d. After the solution was condensed under reduced pressure to 5 ml, addition of ethanol (20 ml) to the solution precipitated **5a**, which was isolated as a colourless powder by filtration (116 mg, 0.029 mmol, 65%). ^1H NMR (500 MHz, D_2O , TMS in a sealed capillary as an external standard) δ 9.15 (d, $J = 6.5$ Hz, 12 H), 9.00 (d, $J = 5.9$ Hz, 12 H), 8.04 (d, $J = 6.5$ Hz, 12 H), 7.78 (d, $J = 5.9$ Hz, 12 H), 7.11 (d, $J = 7.9$ Hz, 12 H), 6.07 (d, $J = 7.9$ Hz, 12 H), 4.40 (s, 12 H), 2.8–2.9 (m, 24 H); ^{13}C NMR (125 MHz, D_2O , TMS in a sealed capillary as an external standard) δ 169.2 (Cq), 169.0 (Cq), 158.2 (Cq), 153.6 (CH), 152.9 (CH), 146.9 (Cq), 144.6 (Cq), 132.5 (Cq), 129.4 (CH), 129.3 (CH), 127.8 (CH), 126.1 (CH), 48.4 (CH_2), 48.3 (CH_2), 42.2 (CH_2). Infrared (KBr) 1,637, 1,523, 1,384, 1,052, 810 cm^{-1} . Mass spectrometry was performed after **5a** was converted its PF_6^- salt by adding NH_4PF_6 to an aqueous solution of **5a**. Exact ESI-MS (PF_6^- salt): prominent peaks for $[\text{M}-(\text{PF}_6)_n]^{n+} \cdot (\text{CH}_3\text{CN})_m$ ($n = 4-9$, $m = 2-18$) were clearly observed: for example, m/z found for $[\text{M}-(\text{PF}_6)_8 + 14\text{CH}_3\text{CN}]^{8+}$, 559.3928; calculated 559.3953.

Crystallographic study of 5a. Slow evaporation of water from an aqueous solution (1.4 ml) of **5a** (20 mg) containing 0.2M NaBF_4 at ambient temperature for 17 d gave a single crystal of **5a** (BF_4^- salt) suitable for X-ray crystallographic analysis. Crystal data: monoclinic, $C2/c$ (no. 15); $a = 30.83$ (1), $b = 22.67$ (1),

$c = 25.731$ (5) Å, $\beta = 109.00$ (4), $V = 17005$ (9) Å³; $z = 8$; $F(000) = 8880$; $\mu(\text{Mo K}\alpha) = 50.14$ cm^{-1} ; $\lambda = 0.71070$; temp, -150°C ; 17,770 reflections measured (10,686 independent), 6,176 observed ($I > 4.00\sigma(I)$); $R = 0.119$; $R_w = 0.153$. In this refinement, twelve counterions and twelve water oxygens were taken into account, including definitely assigned eight counterions (four BF_4^- ions in an asymmetric unit) and four (two) BF_4^- fragments. Hydrogen atoms are located at the calculated positions (not refined). Despite the data collection at -150°C , the quality of the data is inferior owing to the high degree of disorder of BF_4^- anions and solvent molecules. However, all heavy atoms involved in the cage framework are definitely identified.

Self-assembly of 5b. Ligand **3** (54 mg, 0.17 mmol) was added to an aqueous solution (26 ml) of **2b** (75 mg, 0.26 mmol) and the solution was stirred at 80°C for 3 d. On the other hand, a methanol (13 ml) solution of **4** (100 mg, 0.17 mmol) was added to an aqueous solution (1 ml) of **2b** (75 mg, 0.26 mmol). These two solutions were combined and the mixture was stirred at 80°C for 5 d. The addition of ethanol (600 ml) to this solution and stirring the solution at 4°C for 1 d precipitated **5b**, which was isolated as colourless crystals (186 mg, 0.053 mmol, 61%) after filtration. Melting point $>300^\circ\text{C}$; ^1H NMR (500 MHz, D_2O) δ 9.11 (d, $J = 7.1$ Hz, 12 H), 8.96 (d, $J = 6.6$ Hz, 12 H), 8.04 (d, $J = 7.1$ Hz, 12 H), 7.79 (d, $J = 6.6$ Hz, 12 H), 7.07 (d, $J = 8.1$ Hz, 12 H), 6.03 (d, $J = 8.1$ Hz, 12 H), 4.38 (s, 12 H); ^{13}C NMR (125 MHz, D_2O) δ 169.3 (Cq), 169.0 (Cq), 158.3 (Cq), 152.6 (CH), 152.0 (CH), 146.9 (Cq), 144.9 (Cq), 132.5 (Cq), 129.4 (CH), 128.2 (CH), 127.4 (CH), 125.7 (CH), 47.6 (CH_2), 47.5 (CH_2), 42.3 (CH_2); infrared (KBr) 1,637, 1,507, 1,385, 1,059, 826, 806 cm^{-1} . Analysis: calculated for $\text{C}_{126}\text{H}_{132}\text{N}_{48}\text{O}_{36}\text{Pd}_6 \cdot 12\text{H}_2\text{O}$; C, 40.36; H, 4.19; N, 17.93. found: C, 40.16; H, 3.91; N, 17.65.

Received 5 October 1998; accepted 20 April 1999.

1. Wasserman, E. The preparation of interlocking rings: a catenane. *J. Am. Chem. Soc.* **82**, 4433–4434 (1960).
2. Amabilino, D. B. & Stoddart, J. F. Interlocked and intertwined structures and superstructures. *Chem. Rev.* **95**, 2725–2728 (1995).
3. Dietrich-Buchecker, C. O., Sauvage, J.-P. & Kintzinger, J.-P. Une nouvelle famille de molecules: les metallo-catenanes. *Tetrahedr. Lett.* **24**, 5095–5098 (1983).
4. Sauvage, J.-P., Dietrich-Buchecker, C. O. & Chambrion, J.-C. in *Comprehensive Supramolecular Chemistry* Vol. 2 (eds Sauvage, J.-P. & Hosseini, M. W.) Ch. 2, 43–83 (Pergamon, Oxford, 1995).
5. Ashton, P. R. et al. [2]catenane made to order. *Angew. Chem. Int. Edn Engl.* **28**, 1396–1399 (1989).
6. Hunter, C. A. Synthesis and structure elucidation of a new [2]-catenane. *J. Am. Chem. Soc.* **114**, 5303–5311 (1992).
7. Vögtle, F., Meier, S. & Hoss, R. One-step synthesis of a fourfold functionalized catenane. *Angew. Chem. Int. Edn Engl.* **31**, 1619–1621 (1992).
8. Fujita, M., Ibukuro, F., Hagihara, H. & Ogura, K. Quantitative self-assembly of a [2] catenane from two preformed molecular rings. *Nature* **367**, 720–723 (1994).
9. Johnston, A. G., Leigh, D. A., Pritchard, R. J. & Deegan, M. D. Facile synthesis and solid-state structure of a benzylic amide [2] catenane. *Angew. Chem. Int. Edn Engl.* **34**, 1209–1212 (1995).
10. Hamilton, D. G. et al. Tandem hetero-catenation: templating and self-assembly in the mutual closure of two different interlocking rings. *J. Am. Chem. Soc.* **120**, 1096–1097 (1998).
11. Batten, S. R. & Robson, R. Interpenetrating nets: ordered, periodic entanglement. *Angew. Chem. Int. Edn Engl.* **37**, 1460–1494 (1998).
12. Drain, C. M., Nifatis, F., Vasenko, A. & Batteas, J. D. Porphyrin tessellation by design: metal-mediated self-assembly of large arrays and tapes. *Angew. Chem. Int. Edn Engl.* **37**, 2344–2347 (1998).
13. Baxter, P. N. W. in *Comprehensive Supramolecular Chemistry* Vol. 9 (eds Sauvage, J.-P. & Hosseini, M. W.) Ch. 5, 165–211 (Pergamon, Oxford, 1996).
14. Fujita, M. Self-assembly of [2]catenanes containing metals in their backbones. *Acc. Chem. Res.* **32**, 53–61 (1999).
15. Fujita, M. Metal-directed self-assembly of two- and three-dimensional synthetic receptors. *Chem. Soc. Rev.* **27**, 417–425 (1998).
16. Dietrich-Buchecker, C., Frommberger, B., Lürer, I., Sauvage, J.-P. & Vögtle, F. Multiring catenanes with a macrobicyclic core. *Angew. Chem. Int. Edn Engl.* **32**, 1434–1437 (1993).
17. Ashton, P. R. et al. An interwoven supramolecular cage. *Angew. Chem. Int. Edn Engl.* **36**, 59–62 (1997).
18. Fujita, M., Aoyagi, M., Ibukuro, F., Ogura, K. & Yamaguchi, K. Made-to-order assembling of [2]catenanes from palladium(II)-linked rectangular molecular boxes. *J. Am. Chem. Soc.* **120**, 611–612 (1998).
19. Ibukuro, F., Kusukawa, T. & Fujita, M. A thermally switchable molecular lock. Guest-templated synthesis of a kinetically stable nano-sized cage. *J. Am. Chem. Soc.* **120**, 8561–8562 (1998).
20. Jorgensen, W. L. & Severance, D. L. Aromatic-aromatic interactions: free energy profiles for the benzene dimer in water, chloroform, and benzene. *J. Am. Chem. Soc.* **112**, 4768–4774 (1990).
21. Hunter, C. A. & Sanders, L. K. M. The nature of π - π interactions. *J. Am. Chem. Soc.* **112**, 5525–5534 (1990).
22. Fujita, M. et al. Self-assembly of ten molecules into nanometre-sized organic host frameworks. *Nature* **378**, 469–471 (1995).
23. Hasenknopf, B., Lehn, J.-M., Kneisel, B. O., Baum, G. & Fenske, D. Self-assembly of a circular double helicate. *Angew. Chem. Int. Edn Engl.* **35**, 1838–1840 (1996).
24. Fujita, M., Nagao, S. & Ogura, K. Guest-induced organization of a three-dimensional palladium (II) cage-like complex, a prototype for “induced-fit” molecular recognition. *J. Am. Chem. Soc.* **117**, 1649–1650 (1995).

Supplementary information is available on Nature's World-Wide Web site (<http://www.nature.com>) or as paper copy from the London editorial office of Nature.

Correspondence and requests for materials should be addressed to M.F. (e-mail: mfujita@apchem.nagoya-u.ac.jp)

Rapid accumulation and turnover of soil carbon in a re-establishing forest

Daniel D. Richter*, Daniel Markewitz†, Susan E. Trumbore‡ & Carol G. Wells§

* Nicholas School of the Environment, Duke University, Durham, North Carolina 27708, USA

† Warnell School of Forest Resources, University of Georgia, Athens, Georgia 30602, USA

‡ Department of Earth System Science, University of California, Irvine, California 92697, USA

§ US Department of Agriculture, Forest Service, Forest Sciences Laboratory, Research Triangle Park, North Carolina 27709, USA

Present understanding of the global carbon cycle is limited by uncertainty over soil-carbon dynamics^{1–6}. The clearing of the world's forests, mainly for agricultural uses, releases large amounts of carbon to the atmosphere (up to 2×10^{15} g yr⁻¹), much of which arises from the cultivation driving an accelerated decomposition of soil organic matter^{1–4}. Although the effects of cultivation on soil carbon are well studied, studies of soil-carbon recovery after cultivation are limited^{4–11}. Here we present a four-decade-long field study of carbon accumulation by pine ecosystems established on previously cultivated soils in South Carolina, USA⁷. Newly accumulated carbon is tracked by its distinctive ¹⁴C signature, acquired around the onset of forest growth from thermonuclear bomb testing that nearly doubled atmospheric ¹⁴CO₂ in the 1960s. Field data combined with model simulations indicate that the young aggrading forest rapidly incorporated bomb radiocarbon into the forest floor and the upper 60 cm of underlying mineral soil. By the 1990s, however, carbon accumulated only in forest biomass, forest floor, and the upper 7.5 cm of the mineral soil. Although the forest was a strong carbon sink, trees accounted for about 80%, the forest floor 20%, and mineral soil <1%, of the carbon accretion. Despite high carbon inputs to the mineral soil, carbon sequestration was limited by rapid decomposition, facilitated by the coarse soil texture and low-activity clay mineralogy.

Our most precise understanding of how land use affects soil carbon comes from decades-long field experiments that directly estimate soil change under controlled land-management regimes^{7–11}. Although long-term soil studies are highly valued^{11–15}, nearly all such studies examine agricultural ecosystems; remarkably few test soil changes associated with reforestation, reclamation or natural plant succession^{7,16}.

One exception is a four-decade-long experiment at the Calhoun Experimental Forest in South Carolina, USA, which has documented changes in soil chemistry during pine-forest development from 1957 to 1997^{7,10,16–19}. Before 1957, the Calhoun soil was cultivated for cotton and other crops for more than a century, practices that depleted soil organic matter and enhanced nutrient availability from fertilization and liming. The upper 35-cm layer of mineral soils at the Calhoun forest is coarse textured (68% sand, 15% clay), which ensures macroporosity and a highly oxidized environment. Clay minerals that are present are low-activity kaolinite with little potential to physically protect organic carbon inputs from microbial attack. Similar soils are found over extensive areas of southeastern North America and the warm temperate zone and tropics^{20,21}.

Soil studies conducted near the Calhoun forest suggest that from the early 1800s to the middle of the twentieth century, agricultural use (primarily for cotton, corn, hay and pasture) reduced organic carbon in the upper 30 cm of the Calhoun mineral soil by about

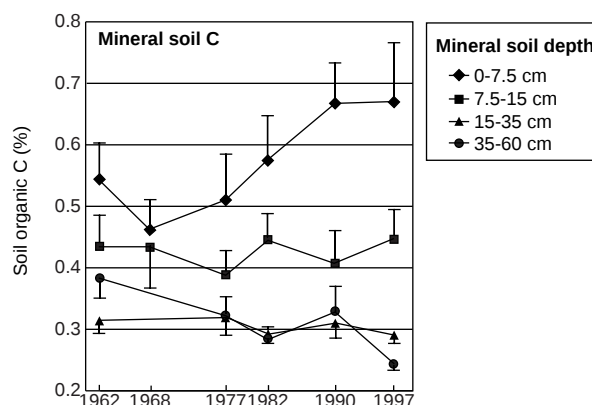


Figure 1 Mineral-soil carbon (1962–97) in eight permanent plots of the Calhoun Experimental Forest, South Carolina, USA. Plots were planted with loblolly pine seedlings in 1957 following long-term cultivation for cotton and other crops. Error bars depict the spatial variation (as standard errors) among the eight permanent plots. The randomized complete block analysis of variance indicated highly significant increases over the 40 years in soil carbon at 0–7.5 cm depth ($p < 0.001$), and decreases at 35–60 cm depth ($p < 0.05$).

1,300 g m⁻², or by ~40% of the carbon present at 0–30-cm depth before cultivation⁷. Here we examine recovery of soil carbon during four decades of reforestation, by using eight permanent plots that were last harvested of cotton in 1955 and were planted in 1957 with seedlings of loblolly pine (*Pinus taeda* L.). On these permanent plots, soil samples have been collected on seven occasions between 1962 and 1997 using the same sampling procedures. Nearly all soil samples are archived and available for analysis.

Since 1957, the aggrading forest accumulated ~3,925 g m⁻² of new carbon in the soil profile—that is, the forest floor (O horizon) plus mineral soil: this represents carbon accumulation of nearly 100 g m⁻² yr⁻¹ for 40 years (Table 1). Most of the new soil carbon, $3,780 \pm 251$ g m⁻² (where ± 251 indicates the standard error), is contained in the forest floor that now blankets the formerly cultivated mineral soil. Only 145 ± 26 g m⁻² of carbon accumulated in the underlying mineral soil (Table 1), about 4% of the new soil carbon. Although carbon in 0–7.5-cm mineral soil increased significantly ($P < 0.001$) between 1962 and the 1990s, there were no significant increases in soil carbon deeper than 7.5 cm (Fig. 1). In fact, carbon in the lowermost soil layer sampled, at 35–60 cm depth, significantly decreased during the 40-year study, perhaps due to slow oxidation of organic-carbon input from crop roots during the period of farming.

Because the Calhoun forest was planted in 1957, new soil carbon derived from the forest has a distinctive isotopic composition because above-ground nuclear-bomb testing in the 1950s and 1960s greatly increased ¹⁴CO₂ in the atmosphere. The fate of new forest carbon can be examined in the ecosystem because the Calhoun forest has grown entirely within the period when the concentration of ¹⁴C has been elevated in both the global atmosphere and in inputs of carbon to the reforested soil.

Changes in soil radiocarbon (¹⁴C) during forest development show that soil carbon has been more dynamic than might be suggested by the gradual reaccumulation of total carbon alone. A decomposition model²² indicates that by 1965, $\Delta^{14}\text{C}$ in the forest floor approached +700‰ only one year after the ¹⁴C peak in the atmosphere (Fig. 2; $\Delta^{14}\text{C}$ is defined in Methods). By the 1990s, forest-floor $\Delta^{14}\text{C}$ declined to less than +300‰, lagging the decrease in atmospheric ¹⁴CO₂ owing to incorporation of ¹⁴C into slow-to-decompose, humic compounds of the acidic pine forest floor (Fig. 2).

By the 1990s, bomb-produced ¹⁴C was most concentrated in the basal layers of the forest floor and the 0–7.5-cm mineral soil (Fig. 2). In 1992, Oe and Oa horizons (the middle and the lowest layers

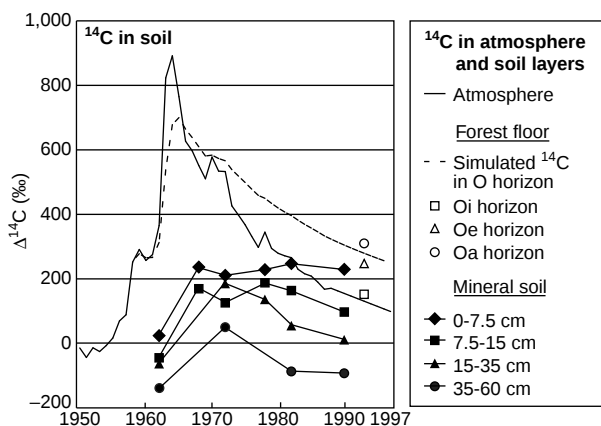


Figure 2 Time trends of ^{14}C in the Calhoun forest ecosystem. Shown are data for atmospheric CO_2 (1950–97); forest floor of Oi (L), Oe (F), and Oa (H) layers (in 1992); and mineral soil (in 1962, 1968, 1972, 1977, 1982 and 1990). Simulated changes in ^{14}C in forest floor (1957–96) are estimated from the decomposition model of Jorgensen and Wells²⁰ and estimates of annual litterfall input over the four decades. The $\Delta^{14}\text{C}$ for the 35–60-cm sample in 1962 was not measured, but was estimated conservatively by a soil-depth-based regression.

of forest floor) had $\Delta^{14}\text{C}$ values of +247.3‰ and +309.8‰, respectively. These older forest litter materials are enriched in ^{14}C derived from plant biomass synthesized during the era of elevated $^{14}\text{CO}_2$ (Fig. 2). The surficial Oi horizon (litterfall deposited within the past three to four years) had $\Delta^{14}\text{C}$ of +152.2‰ in 1992, closely comparable to that of atmospheric CO_2 during the late 1980s and early 1990s (Fig. 2).

Despite the relatively modest 40-year changes in mineral-soil carbon (Table 1), organic matter had incorporated ^{14}C throughout the entire 0–60-cm layer of mineral soil within 8 years after the atmosphere peaked in $^{14}\text{CO}_2$ (Fig. 2). By 1968, $\Delta^{14}\text{C}$ of mineral soil at 0–15-cm depth had increased to +200‰, up from –10.4‰ in 1962. By 1972, $\Delta^{14}\text{C}$ of the entire 0–60-cm mineral soil averaged +125‰, compared to less than –100‰ in 1962.

Since the 1960s, however, only mineral soil at 0–7.5-cm depth has maintained its elevated ^{14}C (Fig. 2). This surficial layer of mineral soil is accumulating carbon in an incipient A horizon that is slowly reforming under the forest following long-term cultivation (Fig. 1). But below 7.5-cm depth, decreases in ^{14}C generally parallel atmospheric decreases in ^{14}C , a pattern that indicates that forest inputs of carbon are being rapidly decomposed (Fig. 2). We suggest that the 0–7.5-cm mineral soil continues to receive ^{14}C from relatively recalcitrant and enriched ^{14}C compounds that mainly reside in the lowest layers of the forest floor (Fig. 2).

To understand better these dynamics, we estimated carbon inputs to soils in the 1990s from three main processes: canopy litterfall, rhizo-deposition (fine-root sloughing and turnover), and hydrological leaching of dissolved organic carbon (DOC) from several sources.

Inputs of carbon to the forest floor in the mid-1990s totalled $\sim 290 \text{ g m}^{-2} \text{ yr}^{-1}$ (Table 1), most of which was from canopy litterfall, although smaller amounts were derived from turnover of fine roots²³ and the DOC in canopy throughfall. The forest's 40-year accretion of forest-floor carbon, $3,780 \text{ g m}^{-2}$, is about 13 times greater than current annual input (Table 1). Accumulation of forest-floor carbon is attributed mainly to complex, acidic, and recalcitrant compounds derived from the coniferous leaf litter. The forest floor is classified as an acidic mor, perched atop the mineral soil, with relatively minimal mixing by soil animals of solid organic material with mineral soil below⁷.

Inputs of carbon to the 0–15-cm layer of mineral soil totalled $\sim 95 \text{ g m}^{-2} \text{ yr}^{-1}$, much of which was from rhizo-deposition, although a substantial fraction was derived from DOC in leachates from the

Table 1 Soil carbon inputs and 40-yr accretions

	Forest floor	0–15 cm	15–60 cm
Annual soil carbon input* ($\text{g m}^{-2} \text{ yr}^{-1}$)			
Annual canopy litterfall†	245 (11.9)	0	0
Annual DOC input‡	8 (13.1)	32 (28.1)	19 (34.5)
Annual rhizo-deposition‡	37 (23.0)	63 (14.1)	26 (26.5)
Annual carbon input	290	95	45
Total soil carbon accretion*† (g m^{-2})	3,780 (18.8)	145 (50.5)	0

* Annual soil carbon inputs (in the 1990s) and four-decade soil carbon accretion in eight permanent plots of the Calhoun Experimental Forest, South Carolina, USA. Reported are means and (in parentheses) coefficients of variation in % among the eight permanent plots. † Soil carbon accretion is estimated for forest floor over 1957–97, and for 0–60-cm mineral soil over 1962–97. Carbon accretions were statistically significant in 0–7.5-cm mineral soils, but not in 7.5–60-cm layers. Bulk densities for mineral soils are 1.52 Mg m^{-3} in 0–35-cm layers of mineral soil and 1.44 Mg m^{-3} at 35–60 cm (ref. 18).

‡ Inputs of canopy litterfall to forest floor were estimated in 1991–92. Inputs of dissolved organic carbon (DOC) were estimated in bi- or tri-weekly collections over two years (1992–94): into forest floor from DOC in canopy throughfall; into 0–15-cm mineral soil from DOC in water infiltrating from the forest floor; and into 15–60-cm soil from DOC in water draining into soil at 15-cm depth. Inputs of carbon from rhizo-deposition were estimated in 1994–95 from 50% of the live fine-root (<2-mm) biomass in forest floor, or mineral soil at 0–15 and 15–30 cm depth. Fine roots were sampled in 1994–95, every three weeks over 18 months. The 50% factor is used as a conservative estimate of carbon inputs from rhizo-deposition.

forest floor (Table 1). Even with carbon input estimated conservatively (Table 1), long-term soil carbon accumulation (145 g m^{-2}) is only 1.5 times that of estimated annual input in the 1990s. The short residence time for carbon input to mineral soils is notable²⁴. Carbon inputs to surficial mineral soil in fine-root biomass and DOC (Table 1) are readily decomposed, as they have little protection from adsorption to organophilic clays in these surface soils with sandy loam textures^{25,26}.

From an ecosystem perspective, this aggrading forest is a strong carbon sink. Accumulation of carbon is especially pronounced in tree biomass and the forest floor relative to that in mineral soil. During the growth of this forest, trees accumulated $14,060 \text{ g m}^{-2}$ of carbon (between 1957 and 1990), compared with $3,780 \text{ g m}^{-2}$ in the forest floor (1957–97) and 145 g m^{-2} in surficial mineral soil (1962–97). Annual carbon accretions thus averaged about 426, 94 and 4 g m^{-2} in the three ecosystem components, respectively.

The overall pattern of carbon sequestration suggests a low carbon-storage potential for mineral soils compared to that in biomass and forest floor; a similar conclusion has been reached from a review of soil-carbon gains under primary vegetative successions⁶. Yet in comparison to other forested soils recovering from previous cultivation, the mineral soil at the Calhoun forest seems to be relatively slow to recover organic carbon. Previous estimates of mineral-soil carbon gains under forests^{4,11} range from 21 to $55 \text{ g m}^{-2} \text{ yr}^{-1}$, while soils at the Calhoun forest gained carbon at $4.1 \text{ g m}^{-2} \text{ yr}^{-1}$ (Table 1).

Although changes in soil carbon are not easy to estimate, the Calhoun experiment provides an approach for making these estimates: especially useful are its well replicated permanent plots, extensive within-plot composite sampling, and soil archive. The main factors driving relatively low carbon accumulations in Calhoun mineral soils are coarse soil texture, low soil surface area, low-activity clay mineralogy²⁶, and warm and humid climate. A network of long-term soil-ecosystem experiments similar to that at the Calhoun, at sites that encompass a range of controlling variables of soil and ecosystem carbon, would greatly facilitate future modelling and management of the carbon cycle^{11,16}. □

Methods

Research area and field experiment. The Calhoun Experimental Forest is one of the world's longest running experiments in which forest-soil properties are measured periodically in replicated permanent plots with all soil samples archived.

The research area is located at 34.5° N, 82° W. Annual precipitation averages 1,170 mm (1950–87) and temperature 16°C. In the early 1800s, primary deciduous forests at the site were cleared, mainly to grow cotton, and the site was managed for row crops, hay and pasture until the mid-twentieth century^{7,16}. Soils are acidic Ultisols, classified as the Appling series (fine, kaolinitic, thermic Typic Kanhapludults). The Appling soil is a common soil of southeastern North America, and is formed from granitic gneiss, the bedrock from which about half the soils in the southern Piedmont region are derived.

In 1957, 16 permanent plots were installed on two cotton fields at the Calhoun Experimental Forest, eight of which are used in these carbon analyses. The mineral soils of these eight plots were sampled in 1962, 1968, 1972, 1978, 1982, 1990 and 1997. At each collection, each plot was sampled at least 20 times with a 2-cm-diameter punch tube in a systematic random fashion at four soil depths: 0–7.5, 7.5–15, 15–35 and 35–60 cm. Within each plot, the ≥ 20 samples per soil depth were composited, that is, in one sample per depth.

Soil archive, radiation and total carbon. The Duke Soil Archive stores air-dried soil samples at room temperature in sealed glass containers. Total soil carbon was analysed in powdered samples with a Perkin-Elmer CHN combustion instrument. Radiocarbon was measured by accelerator mass spectrometry (AMS) on graphite targets prepared from soil organic matter and is reported as $\Delta^{14}\text{C}$, the per mil deviation of $^{14}\text{C}/^{12}\text{C}$ compared with a decay-corrected oxalic acid standard²⁷. Positive values of $\Delta^{14}\text{C}$ indicate presence of bomb-produced ^{14}C , and negative values indicate predominance of old soil organic matter with ^{14}C that has experienced significant radioactive decay (half-life is 5,730 yr). Radiocarbon was estimated using AMS of composite samples made from soil from the eight plots at each depth. Analysis of variance and means separation tests were used to test effects of time on soil carbon accumulation.

Soil carbon inputs. To estimate carbon inputs to soils, carbon in litterfall, fine roots, and soil water were estimated. Litterfall was sampled in each of the eight permanent plots with five collectors (each 0.72 m² in area) per plot. Canopy litterfall was collected monthly during 1991–92²⁸.

Fine roots were sampled volumetrically using a 6-cm-diameter corer that collected undisturbed cores of O horizon and mineral soil from 0–15 and 15–30 cm depths. Soil cores were taken every three weeks for 18 months in 1994–95. Samples were returned to the laboratory where fine roots (<2-mm) were separated from soil by wet-sieving and hand picking. Live roots were separated from dead, and the former were ashed to estimate carbon contents (taken as half the loss on ignition). Only the live fraction of the fine roots is reported here, and carbon inputs from fine-root turnover were simply estimated as 50% of live fine-root carbon in forest floor, or mineral soil at 0–15 and 15–30 cm depth. This factor (50%) is taken to be a conservative estimate of carbon inputs from rhizo-deposition.

DOC was determined in atmospheric precipitation (wet deposition), canopy throughfall, and solutions draining forest floor and several depths of mineral soils to 60 cm (ref. 19). Wet-only precipitation was collected by an Aerochem Metric sampler. Throughfall was collected in three bulk precipitation collectors in each of 8 plots using 16-cm-diameter funnels that were continuously open. Gravitational lysimeters were used to collect water from below O horizons and at 15-cm depth. Tension lysimeters of porous Teflon-quartz design collected solutions at 60-cm depths. Solutions were collected every two or three weeks over two years, 1992–94, and solutions were estimated for DOC concentration by combustion and infrared analysis, after purging solutions of CO₂ and H₂CO₃ by acidification and sparging with N₂ gas.

Received 8 January; accepted 4 May 1999.

1. Schimel, D. S. Terrestrial ecosystems and the carbon cycle. *Glob. Change Biol.* **1**, 77–91 (1995).
2. Oades, J. M. The retention of organic matter in soils. *Biogeochemistry* **5**, 35–70 (1994).
3. Smith, T. M. et al. The global terrestrial carbon cycle. *Wat. Air Soil Pollut.* **70**, 19–37 (1993).
4. Schlesinger, W. H. *Biogeochemistry* (Academic, Orlando, 1997).
5. Johnson, D. W. & Todd, D. E. in *Management of Carbon Sequestration in Soil* (eds Lal, R., Kimble, J. M., Follett, R. F. & Stewart, B. A.) 351–363 (CRC Press, Boca Raton, Florida, 1998).
6. Schlesinger, W. H. Evidence from chronosequence studies for a low carbon-storage potential of soils. *Nature* **348**, 232–234 (1990).
7. Richter, D. D. & Markewitz, D. in *Evaluation of Soil Organic Matter Models* (eds Powlson, D., Smith, P. & Smith, J.) 397–407 (NATO ASI Ser. Vol. 138, Springer, Berlin, 1996).
8. Jenkinson, D. S. The Rothamsted classical experiments: Are they still of use? *Agron. J.* **83**, 2–10 (1991).
9. Leigh, R. A. & Johnston, A. E. *Long-term Experiments in Agricultural and Ecological Sciences* (CAB International, Wallingford, UK, 1994).
10. Harrison, K. G., Post, W. M. & Richter, D. D. Soil carbon turnover in a recovering temperate forest. *Glob. Biogeochem. Cycles* **9**, 449–454 (1995).

11. Smith, P., Powlson, D. S., Smith, J. U. & Elliott, E. T. (eds) *Evaluation and comparison of soil organic matter models*. *Geoderma* **81**, 1–225 (1997).
12. Likens, G. E. (ed.) *Long-term Studies in Ecology: Approaches and Alternatives* (Springer, New York, 1989).
13. Rissler, P. (ed.) *Long-term Ecological Research: An International Perspective* (SCOPE, Wiley, New York, 1991).
14. Powers, R. & Van Cleve, K. Long-term ecological research in temperate and boreal forest ecosystems. *Agron. J.* **83**, 11–24 (1991).
15. Tilman, D. et al. in *Long-term Experiments in Agricultural and Ecological Sciences* (eds Leigh, R. A. & Johnston, A. E.) 305–319 (CAB International, Wallingford, UK, 1994).
16. Richter, D. D. & Markewitz, D. *The Biogeochemistry of a Forested Soil* (Cambridge Univ. Press, in the press).
17. Wells, C. G. & Jorgensen, J. R. in *Forest Soils and Forest Land Management* (eds Bernier, B. & Winget, C. H.) 137–158 (Laval Univ. Press, Quebec, 1975).
18. Richter, D. D. et al. Soil chemical change during three decades in an old-field loblolly pine (*Pinus taeda* L.) ecosystem. *Ecology* **75**, 1463–1473 (1994).
19. Markewitz, D. et al. Three decades of observed soil acidification in the Calhoun Experimental Forest: Has acid rain made a difference? *Soil Sci. Soc. Am. J.* **62**, 1428–1439 (1998).
20. Richter, D. D. & Babbar, L. I. Soil diversity in the tropics. *Adv. Ecol. Res.* **21**, 315–389 (1991).
21. Buol, S. A., Hole, F. D., McCracken, R. J. & Southard, R. J. *Soil Genesis and Classification* (Iowa State Univ. Press, Ames, 1997).
22. Jorgensen, J., Wells, C. G. & Metz, L. J. Nutrient changes in decomposing loblolly pine forest floor. *Soil Sci. Soc. Am. J.* **44**, 1307–1314 (1980).
23. Vogt, K. et al. Organic matter and nutrient dynamics in forest floors in young and mature *Abies amabilis* stands in western Washington, as affected by fine-root input. *Ecol. Monogr.* **53**, 139–157 (1993).
24. Ruark, G. A. Modeling soil temperature effects on in situ decomposition rates for fine roots of loblolly pine. *Forest Sci.* **39**, 118–129 (1993).
25. Van Lear, D. H., Kapeluck, P. R. & Parker, M. M. in *Carbon Forms and Functions in Forest Soils* (eds McFee, W. W. & Kelly, J. M.) 489–501 (Soil Science Soc. of America, Madison, Wisconsin, 1995).
26. Torn, M. S., Trumbore, S. E., Chadwick, O. A., Vitousek, P. M. & Hendricks, D. M. Mineral control of soil organic carbon storage and turnover. *Nature* **389**, 170–173 (1997).
27. Trumbore, S. A. in *Mass Spectrometry of Soils* (eds Boutton, T. & Yamasaki, S.) 311–340 (Dekker, New York, 1996).
28. Urrego, M. J. B. *Nutrient Accumulation in Biomass and Forest Floor of a 34-year-old Loblolly Pine Plantation*. Thesis, North Carolina State Univ. (1993).

Acknowledgements. We thank P. Heine, N. Bethala, S. Buol, R. Christman, M. Hofmockel, V. Jin, J. Raikes, B. Megonigal, K. O'Neill, W. H. Schlesinger, J. Swanson and S. Zheng for contributions to this work. The Calhoun Experimental Forest was supported by the USDA Forest Service, Duke University, the Ecosystems Program of the NSF, and the National Research Initiative Ecosystems and Soils Programs of the USDA.

Correspondence and requests for materials should be addressed to D.D.R. (e-mail: drichter@duke.edu).

Cretaceous age for the feathered dinosaurs of Liaoning, China

Carl C. Swisher III*, Yuan-qing Wang†, Xiao-lin Wang‡, Xing Xu† & Yuan Wang†

* Berkeley Geochronology Center, 2455 Ridge Road, Berkeley, California 94709, USA

† Institute of Vertebrate Paleontology and Paleoanthropology, Chinese Academy of Sciences, Beijing, 100044, People's Republic of China

‡ Natural History Museum, Changchun University of Science and Technology, 6 Ximengzhu Street, Changchun, 130026, People's Republic of China

The ancient lake beds of the lower part of the Yixian Formation, Liaoning Province, northeastern China, have yielded a wide range of well-preserved fossils: the 'feathered' dinosaurs *Sinosauropteryx*¹, *Protarchaeopteryx* and *Caudipteryx*², the primitive birds *Confuciusornis*³ and *Liaoningornis*⁴, the mammal *Zhangheotherium*⁵ and the reportedly oldest flowering plant, *Archaeofructus*⁶. Equally well preserved in the lake beds are a wide range of fossil plants, insects, bivalves, conchostracans, ostracods, gastropods, fish, salamanders, turtles, lizards, the frog *Callobatrachus*⁷ and the pterosaur *Eosipterus*^{1,8}. This uniquely preserved assemblage of fossils is providing new insight into long-lived controversies over bird–dinosaur relationships^{1,2}, the early diversification of birds^{3,9,10} and the origin and evolution of flowering plants⁶. Despite the importance of this fossil assemblage, estimates of its geological age have varied widely from the Late Jurassic to the Early Cretaceous. Here we present the first ⁴⁰Ar/³⁹Ar dates unambiguously associated with the main fossil horizons of the lower part of the Yixian Formation, and thus, for the first time, provide accurate age calibration of this

important fauna. The results of this dating study indicate that the lower Yixian fossil horizons are not Jurassic but rather are at least 20 Myr younger, placing them within middle Early Cretaceous time.

The bulk of the fossils from the lower part of the Yixian Formation come from a small region surrounding the village of Sihetun, located about 20–25 km south of the city of Beipiao (Fig. 1). Fossils in this area occur primarily at two horizons separated stratigraphically by about 30–40 m (ref. 11). Most of the vertebrate fossils listed above come from an interval of a few metres, designated as Bed 6 in our Fig. 2. The angiosperm *Archaeofructus* and the pterosaur *Eosipterus*, as well as other well preserved fossils, are known from the stratigraphically higher fossiliferous level, here referred to as Bed 8 (Fig. 2). Some confusion surrounds the reported occurrence of the dinosaurs *Protarchaeopteryx* and *Caudipteryx*. These dinosaurs were reported as coming from a 'Chaomidianzi' Formation in the Sihetun area. This formally unpublished formation consists of layers that are also known as the 'Jianshangou' beds and are here considered to be equivalent to Beds 1–9 of the lower part of the Yixian Formation, following the terminology in ref. 11. The designation of a separate formation for these beds has not gained wide acceptance in China.

Palaeontologists focusing on various aspects of the fossil assemblage have considered the fauna to be Late Jurassic^{3,4,12–17}, Late Jurassic to Early Cretaceous^{1,2} or Early to late Early Cretaceous^{18–20} in age. Unfortunately, many of these ages are based on comparisons of freshwater invertebrates from sites that are also poorly dated. The Jurassic age for the *Archaeofructus* site was based on comparisons of fossil insects from sites in Siberia and Kazakhstan¹⁶. However, the ages of these sites are equally poorly known, lacking any independent isotopic age control. Looking at the vertebrate fauna, comparison of *Sinosauropteryx* and *Confuciusornis* with *Compsognathus* and *Archaeopteryx*, respectively, from the Solnhofen Limestone of Europe^{3,21}, led some workers to argue for a Late Jurassic (Tithonian) age or an age near the Jurassic/Cretaceous boundary. However, derived features noted in some of the bird fossils (for example, in comparison with *Archaeopteryx*, *Confuciusornis* lacks teeth, possesses a beak, has a much shorter tail and so on), and the occurrence of the ceratopsian dinosaur *Psittacosaurus*²², have led others to argue for a younger Cretaceous age. A Cretaceous Valanginian age has also been proposed, based on pollen and spores associated with the fossil birds *Sinornis*, *Cathyrornis* and *Chaoyangia* from the overlying Jiufotang Formation¹⁰.

Isotopic dates on volcanics from the Yixian Formation have not helped to resolve the controversy. Proponents of a Jurassic age for

the Lower Yixian Formation refer to a K–Ar date of 137 ± 7 Myr and a Rb–Sr date of 143 ± 4 Myr (refs 23, 24) (these uncertainties are 2 standard deviations (s.d.), all others in this report are 1 s.d.). These workers propose an age of around 135 Myr for the Jurassic/Cretaceous boundary, although most workers now accept an age for the boundary of around 144 Myr (ref. 25). Most of the younger age estimates for a Jurassic/Cretaceous boundary are based on low-temperature glauconite dates which we consider to be unreliable. Although the boundary is still not well calibrated, for the purposes of this study we defer to the discussion within ref. 25 and adopt their age of 144 Myr.

⁴⁰Ar/³⁹Ar dates on a volcanic breccia and diabase from the upper part of the Yixian at Jingangshan have been used to argue for a Cretaceous age for the underlying fossils²⁶. The average age of 121.1 ± 0.2 Myr for these dates suggests correlation with the Cretaceous Aptian. ⁴⁰Ar/³⁹Ar dates have also been obtained from volcanic rocks from the lower part of the Yixian Formation. An andesite from Daxinfangzi has been dated at 122.9 ± 0.3 Myr and a basalt from Beipiao at 121.2 ± 0.3 Myr (ref. 26). However, the relationship of these dated volcanics to fossil-bearing horizons has recently come into question. Additional data indicates that the basalt from Beipiao may be an intrusive sill rather than an interbedded flow within the lake beds of the Yixian Formation. Dated volcanics at Jingangshan and Daxinfangzi are geographically distant from the 'feathered' dinosaur and *Confuciusornis* sites and precise stratigraphic relationships are uncertain and require further investigation.

During stratigraphic studies of the Sihetun area in 1997–98, volcanic layers were found interbedded within the main fossiliferous horizons of Bed 6 (Fig. 2). To try and resolve the age of the fossils of the lower Yixian Formation, two tuffs were collected for ⁴⁰Ar/³⁹Ar dating. Dates on these tuffs are potentially better than previous dates for two reasons. First, given the proximity of the tuffs to the fossiliferous layers, problems with correlation and provenance are obviated. Second, the tuffs are pristine water-lain tephra that contain abundant euhedral sanidine feldspars. Sanidine, because of its argon retentivity and potential reproducibility (as established from previous dating studies on replicate analyses of single crystals), is considered to be one of the most reliable materials for ⁴⁰Ar/³⁹Ar dating^{27,28}.

The tuffs were collected from Bed 6 at two sites, about 4 km apart,

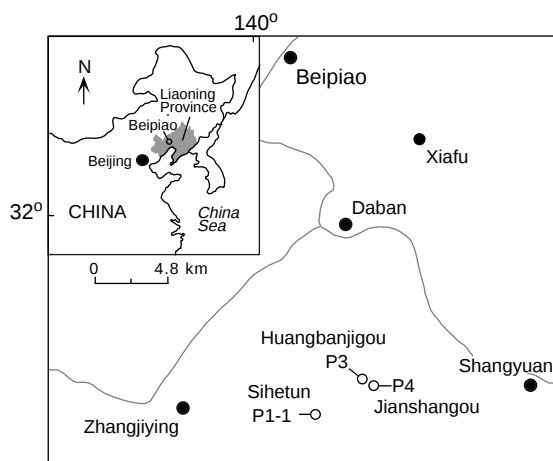


Figure 1 Map of the Sihetun area, Liaoning Province, northeastern China. Locations of stratigraphic sections and sites discussed in the text and in Fig. 2 are shown.

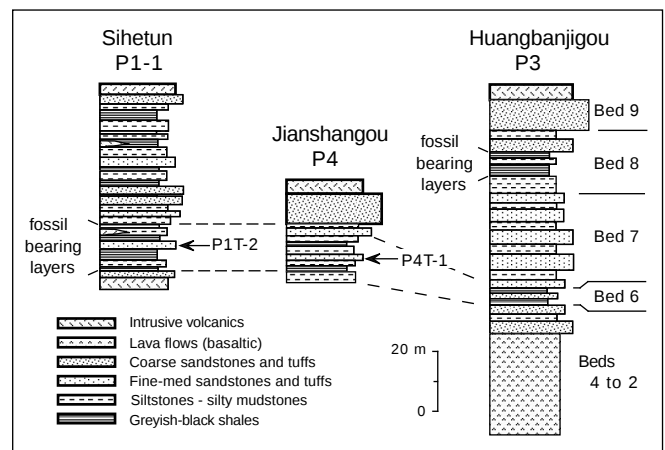


Figure 2 Stratigraphic correlation of the main fossil-bearing layers of the lower Yixian Formation. Arrows show stratigraphic locations of dated tuff samples P1T-2 and P4T-1. Precise stratigraphic relationship of key vertebrate fossils from Bed 6 to the dated tuff layers are reported in the text. The *Archaeofructus* and *Eosipterus*-bearing layers of Bed 8 are located about 30 m stratigraphically above the dated tuff layers. The stratigraphic sections and correlations are modified slightly from those reported in ref. 11 as a result of additional geological field mapping in Autumn of 1998.

in the main fossil-bearing region of Sihetun (Fig. 1, 2). Tuff P4T-1 is from the Jianshangou section (Fig. 2) and occurs about 50 cm above the type specimen of *Zhangheotherium quinquecupidens*^{5,11}. The second tuff (P1T-2) is from the Sihetun section (Fig. 2), located 3.40 m above the layer bearing the type specimen of *Confuciusornis sanctus*^{3,11}. Bed 6 at Sihetun has also yielded well-preserved specimens of *Sinosauropteryx*, *Protarchaeopteryx* and *Liaoningornis*¹¹. The Sihetun and Jianshangou tuffs are stratigraphically within a metre of each other, and may represent the same tuff, although changes in depositional facies preclude this determination. The reportedly oldest flowering plant, *Archaeofructus*, comes from approximately 30 m above this level in Bed 8 at Huangbanjigou (Fig. 1, 2).

⁴⁰Ar/³⁹Ar dating of the P4T-1 and P1T-2 tuffs was done by replicate Ar-ion-laser total-fusion analyses of single crystals of sanidine and CO₂-laser incremental-heating of a bulk sanidine separate. Thirty-four single sanidine crystal dates obtained on Tuff P4T-1 (Fig. 3a) give a mean age of 124.6 ± 0.2 (s.d.) $\pm$ 0.03 Myr (s.e.). Thirty-five single sanidine crystal dates on Tuff P1T-2 (Fig. 3a) give an age of 124.6 ± 0.3 (s.d.) $\pm$ 0.04 Myr (s.e.), an age indistinguishable from that of P4T-1. The radiogenic ⁴⁰Ar yields from all of the sanidines were greater than 99%, indicating little alteration. No crystals of significantly older age that might suggest reworking or detrital contamination were found in either tuff. A few additional crystals that were dated gave lower radiogenic yields, but were determined subsequently to be plagioclase crystals based on their measured ³⁷Ar/³⁹Ar (Ca/K) ratios. These lower radiogenic plagioclase analyses gave a wider variance in age distribution (122.5 Myr to 125 Myr) than the sanidine, considered here to be a result of various alteration of the plagioclase. These analyses are omitted from the calculation of the age of the tuffs.

To address the possibility of argon loss and/or trapped excess argon, CO₂-laser incremental-heating of the sanidine of P4T-1 was also undertaken (Fig. 3b). The release profile (Fig. 3b) was relatively flat with only a slight deviation from a plateau occurring at the lowest three temperature increments. All of the 24 increments that formed the plateau yielded over 95% radiogenic ⁴⁰Ar and totalled

more than 90% of the ³⁹Ar released from the sample. The calculated age for the plateau is 124.6 ± 0.1 Myr (Fig. 3b). The uniformly high radiogenic yields of all 24 increments resulted in a small cluster of points when plotted on an inverse (³⁶Ar/⁴⁰Ar vs ³⁹Ar/⁴⁰Ar) isochron, making any fit unreliable. If we include the three lower radiogenic steps that fell outside the plateau in the isochron, an age of 124.6 ± 0.1 Myr (MSWD (mean sum of weighted deviates) = 1.48) is obtained. The ⁴⁰Ar/³⁶Ar intercept of 286.2 ± 10.7 is not distinguishable from the air ratio. The incremental-heating and single-crystal experiments indicate little alteration of the sanidine, no discernible excess argon and no apparent detrital contamination of the sanidine.

The dates reported here are the first ⁴⁰Ar–³⁹Ar analyses on tuffs interbedded directly in the fossiliferous horizons of the lower Yixian Formation. Our ⁴⁰Ar/³⁹Ar ages for Tuffs P4T-1 and P1T-2 from the Bed 6 fossil-bearing levels are about 10% younger than the previously reported K/Ar and Rb/Sr dates^{23,24}. Although we are unaware of the precise stratigraphic and geographic location of these dated samples, the dates are accompanied by large uncertainties and are generally not considered reliable owing to probably alterations and/or contamination. Our sanidine dates, although 2% older, are more consistent with the ⁴⁰Ar/³⁹Ar dates on andesites and basalts of the Yixian Formation reported in ref. 26. However we consider our dates to be a more accurate calibration of the fossils from the lower Yixian Formation. We base this on the superior argon retentivity of sanidine versus the whole rock and biotite^{27,28} used in their dating study²⁶ and on the unambiguous relationships of our dated tuffs to fossiliferous levels. We also note the possibility of some ⁴⁰Ar loss and/or ³⁹Ar recoil during irradiation that accompanies some of the ⁴⁰Ar/³⁹Ar release spectra shown for those dates²⁶, which may indicate that those samples are altered. At least a percentage of the age difference can be attributed to the age of Hb-3gr, the standard used in ref. 26. We can make a second-order comparison of our dates with those in ref. 26 using the interlaboratory standard MMHb which has been intercalibrated with both Hb-3gr²⁹ and our standard Fish Canyon Sanidine²⁷. Compared with our age for Fish Canyon Sanidine, we obtain ages of 122–124 Myr for the dates in ref. 26, bringing them into close agreement with our sanidine dates for tuffs P4T-1 and P1T-2.

The new dates reported here, in conjunction with the results of ref. 26, indicate that the 'feathered' dinosaurs of Liaoning, although primitive in appearance, are not Late Jurassic or even earliest Cretaceous in age. Compared with the geologic timescale of ref. 25, the dates indicate a correlation with the middle Barremian (mid-Early Cretaceous), at least 20 Myr younger than *Archaeopteryx* from the Late Jurassic (Tithonian) Solnhofen Limestone of Europe. Given the similarities already noted with the Solnhofen and Liaoning fossils, it would appear that aspects of the terrestrial fauna were part of a long-lived chronofauna, persisting across the Jurassic–Cretaceous boundary. However, the ages of many of these sites worldwide are poorly known and, in light of the new dates for the lower Yixian fauna reported here, their ages may need re-examination. A final outcome of the new dates is that *Archaeofructus*, although remarkably well preserved, cannot be considered as early as originally thought⁵. Depending on the accuracy in dating of other fossil sites worldwide, *Archaeofructus* appears to be comparable in age with early angiosperm evidence from the Barremian of China, Europe, Russia and eastern North America³⁰. □

Methods

⁴⁰Ar/³⁹Ar dating of the P4T-1 and P1T-2 tuffs follow procedures described in ref. 27 and references therein. Sanidine from the tuffs and multiple samples of the monitor mineral Fish Canyon Sanidine (to determine lateral flux gradients during irradiation) were co-irradiated for 20 h in the cadmium-lined core (CLICIT) facility of the Oregon State University TRIGA reactor. Nucleogenic interference corrections are those previously reported²⁷. Mass discrimination was monitored before and after measurement of the sanidines using an on-line

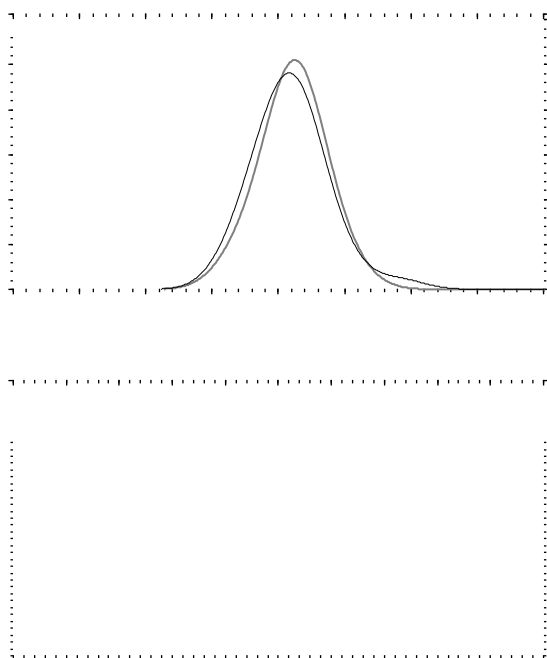


Figure 3 ⁴⁰Ar/³⁹Ar dating results. **a**, Age probability diagram for single sanidine crystal dates; **b**, incremental-heating spectrum on sanidine.

air pipette system. The P4T-1 and P1T-2 dates are referenced to the Fish Canyon Sanidine monitor mineral using the published age of 28.02 Myr (ref. 27). The ages reported are arithmetic means of the individual dates based on total fusion analyses of single sanidine crystals. We report one standard deviation (s.d.) and standard error of the mean (s.e.) uncertainties with the mean ages to document crystal-to-crystal age variation. Individual crystal dates are presented in the Supplementary Information, Table 1.

In the incremental-heating experiment, 15 handpicked euhedral sanidine crystals were laid flat on the bottom of a 2-cm-diameter well of a copper sample disk. The crystals were heated by a CO₂ laser beam directed through an integrator lens that produces relatively even and uniform heating. Twenty-seven discrete temperature increments based on a controlled stepwise increase in power of the CO₂ laser were obtained. The heating time for each increment was 60 s. Incremental heating data are supplied in the Supplementary Information, Table 2.

Received 8 February; accepted 31 March 1999.

- Chen, P.-J., Dong, Z.-M. & Zhen, S.-N. An exceptionally well-preserved theropod dinosaur from the Yixian Formation of China. *Nature* **391**, 147–152 (1998).
- Ji, Q., Currie, P. J., Norell, M. A. & Ji, S.-A. Two feathered dinosaurs from northeastern China. *Nature* **393**, 753–761 (1998).
- Hou, L.-H., Zhou, Z.-H., Martin, L. D. & Feduccia, A. A beaked bird from the Jurassic of China. *Nature* **377**, 616–618 (1995).
- Hou, L.-H. A carnate bird from the Upper Jurassic of western Liaoning, China. *Chinese Sci. Bull.* **41**, 1861–1864 (1996).
- Hu, Y.-M., Wang, Y.-Q., Luo, Z.-X. & Li, C.-K. A new symmetrodont mammal from China and its implications for mammal evolution. *Nature* **390**, 137–142 (1997).
- Sun, G., Dilcher, D. L., Zheng, S.-L. & Zhou, Z.-K. In search of the first flower: A Jurassic angiosperm, *Archaeofructus*, from northeast China. *Science* **282**, 1692–1695 (1998).
- Wang, Y. & Gao, K.-Q. Earliest Asian discoglossid frog from western Liaoning. *Chinese Sci. Bull.* **44**, 636–642 (1999).
- Ji, S.-A. & Ji, Q. Discovery of a new pterosaur from western Liaoning, China. *Acta Geol. Sin.* **71**, 1–5 (1997).
- Hou, L.-H., Martin, L. D., Zhou, Z.-H. & Feduccia, A. Early adaptive radiation of birds: evidence from fossils from northeastern China. *Science* **274**, 1164–1167 (1996).
- Sereno, P. C. & Rao, C.-G. Early evolution of avian flight and perching: new evidence from the Lower Cretaceous of China. *Science* **255**, 845–848 (1992).
- Wang, X.-L. *et al.* Stratigraphic sequence and vertebrate-bearing beds of the lower part of the Yixian Formation, in Sihetun and neighboring area, western Liaoning, China. *Vert. Palasiat.* **36**, 81–96 (1998).
- Chen, P.-J. Distribution and migration of the Jehol Fauna with reference to nonmarine Jurassic-Cretaceous boundary in China. *Acta Palaeont. Sin.* **27**, 659–683 (1988).
- Wang, S.-E. Origin, evolution and mechanism of the Jehol Fauna. *Acta Geol. Sin.* **63**, 351–360 (1990).
- Yu, J.-S., Dong, G.-Y. & Yao, P.-Y. In *Mesozoic Stratigraphy and Palaeontology of Western Liaoning* Vol. 3, 1–28 (Geological Publishing House, Beijing, 1987).
- Zhang, L.-J. In *Mesozoic Stratigraphy and Palaeontology of Western Liaoning* Vol. 2, 1–212 (Geological Publishing House, Beijing, 1985).
- Ren, D. Flower-associated Brachycera flies as fossil evidence for Jurassic angiosperm origins. *Science* **280**, 85–88 (1998).
- Wang, W.-L. In *Mesozoic Stratigraphy and Palaeontology of Western Liaoning* Vol. 3, 134–201 (Geological Publishing House, Beijing, 1987).
- Hao, Y.-C., Su, D.-Y., Li, Y.-G. & Li, P.-X. Stratigraphical division of non-marine Cretaceous and the Juro-Cretaceous boundary in China. *Acta Geol. Sin.* **56**, 187–199 (1982).
- Li, P.-X., Su, D.-Y., Li, Y.-G. & Yu, J.-X. The chronostratigraphic status of the *Lycoptera*-bearing bed. *Acta Geol. Sin.* **68**, 87–100 (1994).
- Mao, S., Yu, J. & Lentini, J. K. Palynological interpretation of Early Cretaceous non-marine strata of northeast China. *Rev. Paleobot. Palynol.* **65**, 115–118 (1990).
- Wellnhofer, P. Das siebte Exemplar von *Archaeopteryx* aus den Solnhofener Schichten. *Archaeopteryx* **11**, 1–48 (1993).
- Xu, X. & Wang, X.-L. New pterosaur (Ornithischia, Ceratopsia) occurrence from the Yixian Formation of Liaoning, China, and its stratigraphical significance. *Vert. Palasiat.* **36**, 147–158 (1997).
- Wang, D.-F. On the age of the Rehe Group in western Liaoning Province, China. *Bull. Chinese Acad. Geol. Sci.* **7**, 65–82 (1983).
- Wang, D.-F. & Diao, N.-C. in *Scientific Papers on Geology for International Exchange* Vol. 5, 1–12 (Geological Publishing House, Beijing, 1984).
- Gradstein, F. M. *et al.* in *Geochronology, Time Scales and Global Stratigraphic Correlation* (eds Berggren, W. A., Kent, D. V., Aubry, M.-P. & Hardenbol, J.) 95–126 (S.E.P.M. Spec. Publ. No. 54, 1995).
- Smith, P. *et al.* Dates and rates in ancient lakes: ⁴⁰Ar/³⁹Ar evidence for an Early Cretaceous age for the Jehol Group, northeast China. *Can. J. Earth Sci.* **32**, 1426–1431 (1995).
- Renne, P. R. *et al.* Intercalibration of standards, absolute ages and uncertainties in ⁴⁰Ar/³⁹Ar dating. *Chem. Geol.* **145**, 117–152 (1998).
- McDougall, I. & Harrison, T. M. *Geochronology and Thermochronology* (Oxford Univ. Press, New York, 1988).
- Roddick, J. C. High precision intercalibration of ⁴⁰Ar/³⁹Ar standards. *Geochim. Cosmochim. Acta* **47**, 887–898 (1983).
- Sun, G. & Dilcher, D. L. Early angiosperms from the Lower Cretaceous of Jixi, China and their significance for the study of the earliest occurrence of angiosperms in the world. *Palaeobotany* **45**, 393–399.

Supplementary information is available on Nature's World Wide Web site (<http://www.nature.com>) or as paper copy from the London editorial office of Nature.

Acknowledgements. We thank all the members of the IVPP field crew, and M. M. Chang, Z. Luo and S. Kelley for discussion and comment on this study. This research was funded, in part, by the Chinese Academy of Sciences and the National Science Foundation, Earth Science Program.

Correspondence and requests for materials should be addressed to C.C.S. (e-mail: cswish@bgc.org) or Y.-Q.W. (e-mail: wang.yuanqing@pa.ivpp.ac.cn).

Hydrologically defined niches reveal a basis for species richness in plant communities

Jonathan Silvertown*, Mike E. Dodd*, David J. G. Gowing† & J. Owen Mountford‡

* Ecology and Conservation Research Group, Biology Department, Open University, Walton Hall, Milton Keynes MK7 6AA, UK

† Silsoe College, Cranfield University, Silsoe, Bedfordshire MK45 4DT, UK

‡ NERC Institute of Terrestrial Ecology, Monks Wood, Abbot's Ripton, Huntingdon PE17 2LS, UK

Species-rich plant communities are prized repositories of biodiversity and a dwindling resource, but how the large numbers of species that characterize such communities are able to coexist is poorly understood. Resource-based competition theory predicts that stable coexistence between species depends on each being a superior competitor in its own niche¹. The theoretical problem is that plants all require the same resources and acquire them in a very limited variety of ways, so observed niche overlaps are high^{2,3} and exclusion of all but the best competitor is the predicted result. This problem, central to community ecology, has elicited a variety of theoretical solutions^{4–7}, several of which invoke some degree of niche separation in time or space^{8,9}. The signature of niche separation in the field is to be found in community structure, which should indicate (i) smaller than expected niche overlaps on relevant niche axes and (ii) a trade-off between species' resource use on orthogonal axes. Here we provide evidence for the existence of both these conditions in a species-rich plant community.

We sampled two English meadow plant communities (flood-plain meadows, NVC MG8 and MG4 (ref. 10)), at Tadham Moor, Somerset, UK, and at Cricklade, Wiltshire, UK. The percentage abundance of all species present was estimated in 844 1-m² quadrats within a 22-ha area at Tadham and in 641 quadrats within a 44-ha area at Cricklade. Species' tolerances were estimated from the range of hydrological conditions in which they were recorded growing in the field. Soil moisture conditions in each quadrat were determined from quadrat locations in relation to surrounding water courses by using two hydrological models parameterized with 15 years of data on meteorological inputs and weekly water levels in the surrounding rivers and ditches. The models were verified against water-table depths measured in dip-wells located at a subset of quadrat locations^{11,12}. Two sum exceedence parameters¹³ were derived from the modelled water-table depths and were used as niche axes. A sum exceedence value (SEV) for soil drying was cumulated during periods in which the moisture tension of the surface soil exceeded 5 kPa, which could potentially induce stomatal closure¹⁴. An SEV for aeration was cumulated during periods in which the soil air-filled porosity fell below 10% by volume, which is assumed to preclude the free diffusion of oxygen in the topsoil¹⁵. High aeration SEVs indicate waterlogging. The water-table depths that, under average summer evaporative demand, gave rise to (i) a surface tension of 5 kPa and (ii) an air-filled porosity of 10% were calculated for each of the soil types. These depths were then used as maximum and minimum thresholds, respectively. Each SEV gives the number of weeks for which the relevant threshold was exceeded multiplied by the vertical distance by which the water-table exceeded it, measured in units of metre weeks. SEVs have the advantage that they incorporate a measure of long-term temporal variation in soil moisture at a scale relevant to the physiological tolerances of plants.

Quadrat sampling locations were distributed randomly throughout the available niche space, as defined by the two SEV axes (Fig. 1). The Tadham site (Fig. 1a) is topographically very flat, whereas

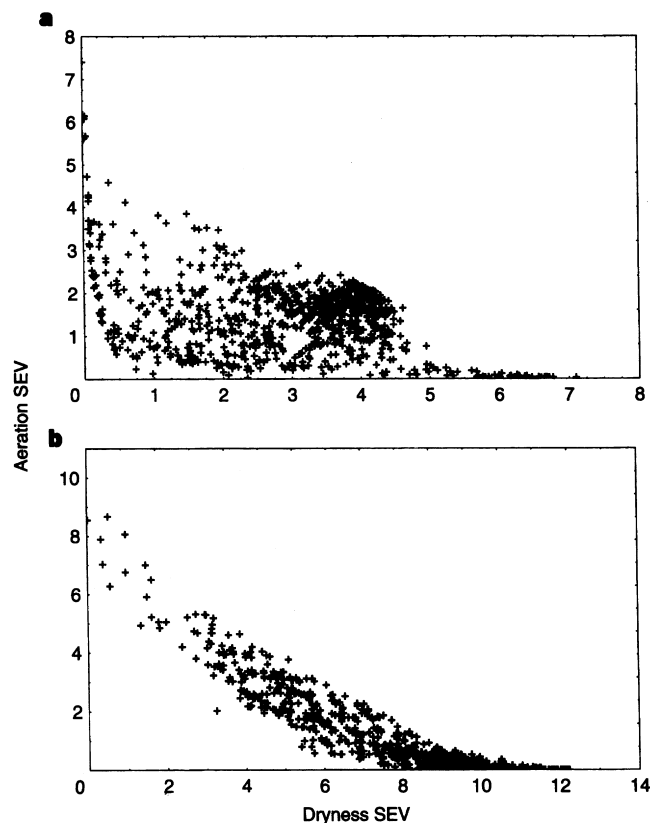


Figure 1 Hydrologically defined niche space. Niche space at Tadham (a) and Cricklade (b). High aeration SEVs indicate frequently waterlogged conditions; high soil-dryness SEVs indicate frequently droughted conditions. Each point indicates a sampling location for plant community composition.

Cricklade has more variable microtopography resulting in a well-defined band of negatively correlated SEV parameters (Fig. 1b).

Pairwise values of Pianka's index of niche overlap¹⁶ were computed for all combinations of species occurring in 50 or more quadrats (Tadham, 64 species; Cricklade, 51 species). Scale-dependence is to be expected in niche overlap, so indices were computed at two scales by dividing niche space at each site into 'boxes' of 0.5 SEV $\times$ 0.5 SEV (fine scale) and 1 SEV $\times$ 1 SEV (coarse scale). We calculated the mean abundance per box for each species, standardized to a total of 100% per species across all boxes. Departures of mean niche overlap for the whole community from random expectation were determined by using a randomization test¹⁷ that randomized the non-zero abundances of species in boxes but kept zero abundances fixed. Mean pairwise niche overlaps were significantly less than expected ($P < 0.0001$) at both fine and coarse scales at Tadham and at a fine scale at Cricklade. At the coarse scale, observed niche overlap at Cricklade was not significantly less than expectation ($P = 0.25$). These results demonstrate that hydrologically defined niches structure both communities.

Experimental support for this conclusion comes from a reanalysis that we have performed of the results of a classic experiment by Ellenberg¹⁸. He sowed six meadow grass species, including four that occurred and two that had close relatives in our study, in single-species stands and in mixture evenly across an experimental water-table gradient spanning 5–140 cm mean water-table depth. SEVs for aeration and soil dryness on the gradient can be estimated; these span the entire range of values observed at Tadham and Cricklade. Dividing Ellenberg's hydrological gradient into ten 'niche boxes' and applying the same randomization test of niche overlap that we used on our own data, we find a degree of niche overlap (median 0.94) among species when grown in monoculture that is far higher than random expectation ($P < 0.0001$). All species showed a peak in

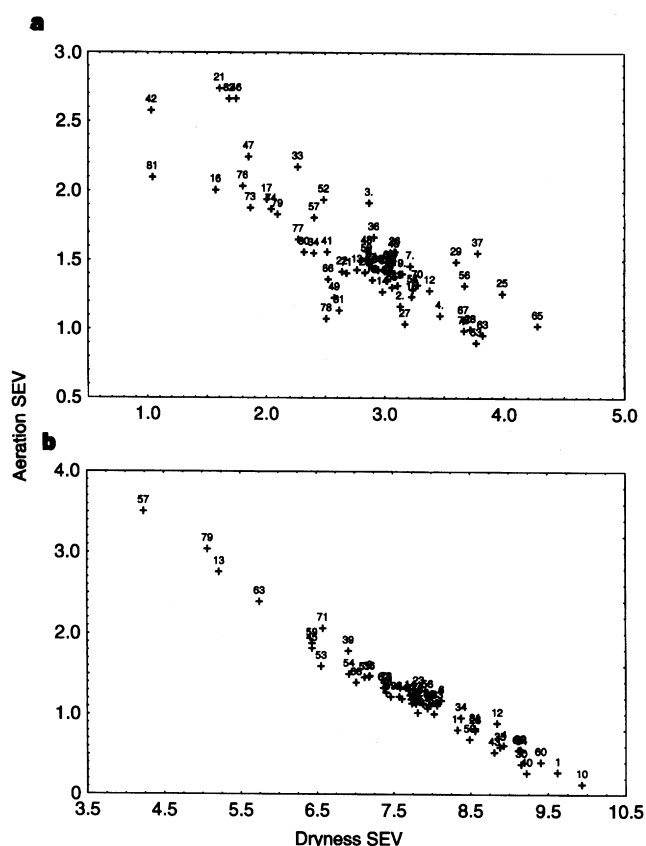


Figure 2 Trade-offs between dryness and aeration SEVs. Mean SEVs weighted by abundance for all species occurring in 50 or more quadrats at Tadham (a) and Cricklade (b). Key to species: 1, *Linum catharticum*; 2, *Filipendula ulmaria*; 3, *Potentilla anserina*; 4, *P. reptans*; 5, *Sanguisorba officinalis*; 6, *Lotus corniculatus*; 7, *L. uliginosus*; 8, *Trifolium repens*; 9, *T. pratense*; 10, *Medicago lupulina*; 11, *Lathyrus pratensis*; 12, *Vicia cracca*; 13, *Cardamine pratensis*; 14, *Cerastium fontanum*; 15, *Stellaria graminea*; 16, *Lychnis flos-cuculi*; 17, *Polygonum amphibium*; 18, *Rumex crispus*; 19, *R. acetosa*; 20, *Myosotis discolor*; 21, *Galium palustre sens. lat.*; 22, *Veronica serpyllifolia*; 23, *Rhinanthus minor*; 24, *Plantago lanceolata*; 25, *Glechoma hederacea*; 26, *Prunella vulgaris*; 27, *Cirsium palustre*; 28, *C. arvense*; 29, *Centaurea nigra* agg.; 30, *Tragopogon pratensis*; 31, *Bellis perennis*; 32, *Leucanthemum vulgare*; 33, *Senecio aquaticus*; 34, *Leontodon hispidus*; 35, *L. taraxacoides*; 36, *L. autumnalis*; 37, *Hypochoeris radicata*; 38, *Taraxacum* agg.; 39, *Silva silaus*; 40, *Heracleum sphondylium*; 41, *Lysimachia nummularia*; 42, *Caltha palustris*; 43, *Ranunculus bulbosus*; 44, *R. acris*; 45, *R. repens*; 46, *Glyceria maxima*; 47, *G. fluitans*; 48, *Bromus racemosus*; 49, *B. hordeaceus sens. lat.*; 50, *B. commutatus*; 51, *Hordeum secalinum*; 52, *Alopecurus geniculatus*; 53, *A. pratensis*; 54, *Anthoxanthum odoratum*; 55, *Agrostis canina sens. lat.*; 56, *A. capillaris*; 57, *A. stolonifera*; 58, *Arrhenatherum elatius*; 59, *Phleum pratense subsp. pratense*; 60, *Trisetum flavescens*; 61, *Holcus lanatus*; 62, *Cynosurus cristatus*; 63, *Deschampsia cespitosa subsp. cespitosa*; 64, *Briza media*; 65, *Dactylis glomerata*; 66, *Poa trivialis*; 67, *P. subcaerulea*; 68, *P. pratensis sens. lat.*; 69, *Lolium perenne subsp. perenne*; 70, *Festuca rubra* agg.; 71, *F. pratensis*; 72, *Juncus articulatus*; 73, *J. effusus*; 74, *J. inflexus*; 75, *Luzula campestris*; 76, *Carex disticha*; 77, *C. flacca*; 78, *C. hirta*; 79, *C. nigra*; 80, *C. panicea*; 81, *C. riparia*; 82, *Eleocharis palustris*; 83, *Fritillaria meleagris*.

biomass production in monoculture at between 20 and 35 cm design water-table depth. In contrast, niche overlap among the same species when in competition with one another (as in the field) was significantly reduced ($P < 0.0001$; median overlap 0.72); species segregated along the water-table gradient, showing peaks in biomass over a wider range of 5–110 cm depending on species. This reanalysis not only demonstrates that our field observations can be reproduced experimentally, but also indicates that the null model of random niche overlap that we have used to detect the community

structure generated by hydrological conditions is very conservative because plants grown in the absence of interspecific competition have highly overlapping rather than randomly overlapping fundamental niches.

We used our field data to test for a trade-off between SEV axes. A mean SEV for each species on each axis at each site was calculated as $(\sum_{i=1}^n \text{SEV}_i p_i)/n$, where n was the number of quadrats in which a species was present, SEV_i was the value at the location of quadrat i , and p_i was the proportion of the species' total recorded abundance found in quadrat i . The distribution of mean SEV values in niche space (Fig. 2) shows a highly significant negative correlation between soil drying and waterlogging tolerance across species at both sites (Tadham, $r = 0.82$, $n = 64$, $P < 0.0001$ (Fig. 2a); Cricklade, $r = 0.98$, $n = 51$, $P < 0.0001$ (Fig. 2b)), indicating a trade-off between tolerances.

We used the method of phylogenetically independent contrasts, as implemented in the computer program CAIC¹⁹, to check the robustness of the trade-off between drought and waterlogging tolerance (Fig. 2). A composite phylogeny (available from the authors), with branch lengths set equal, was constructed for the 83 species in the study by using the most up-to-date molecular phylogenies available. Regression through the origin of aeration and dryness SEV contrasts gave a highly significant negative relationship at both sites (Tadham, $y = -1.20x$, $F_{(1,52)} = 87.12$, $P < 0.0001$, adjusted $R^2 = 0.62$; Cricklade, $y = -1.68x$, $F_{(1,44)} = 1,279$, $P < 0.0001$, adjusted $R^2 = 0.966$). The trade-off between soil-drying and waterlogging tolerance is therefore not only strong, but has shaped the evolution of tolerances in many independent cases²⁰.

For at least two reasons, meadows are unlikely to be the only species-rich communities that are structured by segregation of species along niche axes of soil drying and soil aeration. First, segregation across topographic gradients has been observed in many plant communities, and we have shown that the spatial variation in hydrological conditions thought to cause this might occur in the absence of any obvious topographic variation (as at Tadham) and that plants are sensitive to hydrology at a fine scale. Second, the trade-off between drought and aeration tolerance that we have demonstrated here is so strong and phylogenetically robust that it cannot be confined to our set of species alone and must reflect a general physiological constraint with wide occurrence. The hitherto largely vain search for plant niches in the established phase of the plant life cycle has been confined mainly to axes of nutrients and light availability, but our findings suggest that investigations of variation in soil hydrology at a fine spatial scale and over longer time-spans might well reveal a potent force that structures many types of plant community. The precise mechanism by which this force acts now requires investigation. □

Received 15 February; accepted 19 May 1999.

- Williamson, M. H. An elementary theory of interspecific competition. *Nature* **180**, 422–425 (1957).
- Mahdi, R., Law, R. & Willis, A. J. Large niche overlaps among coexisting plant species in a limestone grassland community. *J. Ecol.* **77**, 386–400 (1989).
- Tilman, G. D. Plant dominance along an experimental nutrient gradient. *Ecology* **65**, 1445–1453 (1984).
- Chesson, P. L. Coexistence of competitors in spatially and temporally varying environments: a look at the combined effects of different sorts of variability. *Theor. Popul. Biol.* **28**, 263–287 (1985).
- Hurt, G. C. & Pacala, S. W. The consequences of recruitment limitation—reconciling chance, history and competitive differences between plants. *J. Theor. Biol.* **176**, 1–12 (1995).
- Hubbell, S. P. & Foster, R. B. In *Community Ecology* (eds Diamond, J. & Case, T. J.) 314–329 (Harper Row, New York, 1986).
- Tilman, D. *Resource Competition and Community Structure* (Princeton Univ. Press, 1982).
- Shmida, A. & Ellner, S. P. Coexistence of plants with similar niches. *Vegetatio* **58**, 29–55 (1984).
- Pacala, S. W. in *Plant Ecology* (ed. Crawley, M. J.) 532–555 (Blackwell, Oxford, 1997).
- Rodwell, J. S. *British Plant Communities* Vol. 3 (Cambridge Univ. Press, 1992).
- Youngs, E. G., Leeds-Harrison, P. B. & Chapman, J. M. Modelling water-table movement in flat low-lying lands. *Hydrol. Process.* **3**, 301–315 (1989).
- Gowing, D. J. G. & Youngs, E. G. *The Effect of the Hydrology of a Thames Flood Meadow on its Vegetation* 68–80 (British Hydrological Society, Occ. Pap. 8, London, 1997).
- Sieben, W. H. Het verband tussen outwatering en onbrengst bij de jonge zavelgronden in de Noordooipolder. *Van Zee Land* **40**, 1–117 (1965).
- Hensen, I. E., Jensen, C. R. & Turner, N. C. Leaf gas exchange and water relations of lupins and wheat. I. Shoot responses to soil water deficits. *Aust. J. Plant Physiol.* **16**, 401–413 (1989).
- Wesseling, J. & van Wijk, W. R. in *Drainage of Agricultural Lands* (ed. Luthin, J. N.) 461–504 (American Society for Agronomy, Madison, Wisconsin, 1957).
- Pianka, E. R. The structure of lizard communities. *Annu. Rev. Ecol. Syst.* **4**, 53–74 (1973).

- Gotelli, N. J. & Entsminger, G. L. *EcoSim ver 1.11: null models software for ecology* (Keseey-Bear, Burlington, 1997).
- Ellenberg, H. Physiologisches und ökologisches Verhalten derselben Pflanzenarten. *Ber. Dtsch. Bot. Ge.* **65**, 350–361 (1953).
- Purvis, A. & Rambaut, A. Comparative analysis by independent contrasts (CAIC): an Apple Macintosh application for analysing comparative data. *Computer Appl. Biosci.* **11**, 247–251 (1995).
- Harvey, P. H. & Pagel, M. D. *The Comparative Method in Evolutionary Biology* (Oxford Univ. Press, 1991).

Acknowledgements. We thank the staff of ITE, IGER and Ecological Surveys (Bangor) for data collection, J. Gilbert for initial data analysis, and G. Spoor and E. Youngs for hydrological interpretation. We acknowledge funding from the Ministry of Agriculture, Fisheries and Food for fieldwork and from the Open University Research Development Committee for data analysis.

Correspondence should be addressed to J.S. (e-mail: j.silvertown@open.ac.uk).

Ectoparasite infestation and sex-biased local recruitment of hosts

Philipp Heeb*, Isabelle Werner*, A. C. Mateman†, Mathias Kölliker*, Martin W. G. Brinkhof*, C. M. Lessells† & Heinz Richner*

* Zoology Department, University of Bern, Centre for Behaviour and Evolution, CH-3032 Hinterkappelen-Bern, Switzerland

† Netherlands Institute of Ecology, Boterhoeksestraat 22, PO Box 40, 6666 ZG Heteren, The Netherlands

Dispersal patterns of organisms are a fundamental aspect of their ecology, modifying the genetic and social structure of local populations^{1–4}. Parasites reduce the reproductive success and survival of hosts and thereby exert selection pressure on host life-history traits^{4–6}, possibly affecting host dispersal^{7–9}. Here we test experimentally whether infestation by hen fleas, *Ceratophyllus gallinae*, affects sex-related recruitment of great tit, *Parus major*, fledglings. Using sex-specific DNA markers, we show that flea infestation led to a higher proportion of male fledglings recruiting in the local population in one year. In infested broods, the proportion of male recruits increased with brood size over a three year period, whereas the proportion of male recruits from uninfested broods decreased with brood size. Natal dispersal distances of recruits from infested nests were shorter than those from uninfested nests¹⁰. To our knowledge, this study provides the first evidence for parasite-mediated host natal dispersal and local recruitment in relation to sex. Current theory needs to consider parasites as potentially important factors shaping life-history traits associated with host dispersal.

As part of a long-term study on the evolutionary and ecological aspects of host–ectoparasite interactions^{11–13}, we randomly infested half the nests in a population of great tits with adult hen fleas¹¹ over three breeding seasons. The remaining nests were kept free of fleas (see Methods). Similar to findings in another Swiss population¹⁴, flea infestation reduced the body mass of nestlings (mean mass ($\pm$ s.e.) of 16-day-old nestlings: infested, 15.19 ± 0.16 g; uninfested, 15.84 ± 0.15 g; two-way ANOVA: flea treatment, $F(1, 283) = 10.1$, $P = 0.002$; year, $F(2, 283) = 26.7$, $P < 0.0001$; interaction, NS). Furthermore, the number of young fledged was lower in flea-infested than in uninfested broods (infested, 4.0 ± 0.2 , $N = 162$ broods; uninfested, 4.7 ± 0.2 , $N = 169$; two-way ANOVA for ranks: fleas, $\chi^2_1 = 7.28$, $P = 0.007$; year, $\chi^2_2 = 4.015$, $P < 0.0001$; interaction, NS). Thus, flea infestation had detrimental effects on great tit reproduction during the nestling period.

During the three years of the experiment, uninfested broods ($N = 150$) fledged 797 young of which 63 (7.9%) were recaptured as first-time breeders (that is, recruits; see Methods). Infested broods ($N = 137$) fledged 654 young of which 48 (7.3%) were recruited. Flea infestation had no significant effect on the total proportion of broods recruiting young (log likelihood ratio: fleas, $\chi^2_1 = 0.34$,

Table 1 Logistic regression on the proportion of fledglings per brood recruited locally

	(Change in) d.f.	(Change in) deviance	P
Null model	280	300.4	
Full model	275	204.4	
Year	2	75.5	< 0.0001
Nestling mass	1	11.8	< 0.001
Year*nestling mass	2	6.0	0.05
Rejected terms			
Flea infestation	1	0.07	0.79
Laying date	1	2.80	0.10
Year*flea infestation	2	0.12	0.90

Logistic regression²⁰ on the number of young recruited with the number of young fledged as the binomial denominator. Degrees of freedom in the null model indicate the number of broods. The effect of fleas remains nonsignificant when nestling mass is excluded from the model ($P > 0.50$).

$P = 0.55$; year, $\chi^2_2 = 57.53$, $P < 0.0001$; interaction, NS), on the number of recruits per brood (Wilcoxon two-sample test: $Z = -0.81$, $P = 0.42$) or on the proportion of fledglings recruited per brood (Table 1). As found in other studies on passerines, the probability of recruitment of fledglings was positively correlated with their mass¹⁵ (Table 1). In our study, flea infestation reduced the body mass of nestlings and the number of fledglings, but there was no evidence that fleas affected overall local recruitment. The relatively low reproductive performance of our great tit population compared with other Swiss populations indicates that our study area is a rather poor habitat^{14,16}. One possible explanation for the absence of an effect of fleas on local recruitment could be that more fledglings from uninfested nests dispersed away from our study area^{17,18}.

It has been suggested that the detrimental effects of parasites on adults depends on the sex of the host¹⁹. To examine whether flea infestation affected the recruitment of fledglings in relation to gender, we used sex-specific DNA-markers to determine the sex of young from 64 broods in one year (see Methods). Results showed that the proportion of males in the brood at fledging did not differ between flea treatments ($Z = 0.14$, $P = 0.9$; Fig. 1), and did not differ significantly from parity (both $P > 0.5$). However, we found that the proportion of male recruits differed in relation to flea infestation ($Z = 2.15$, $P = 0.03$). A greater proportion of male fledglings was recruited from flea-infested broods, whereas recruitment in uninfested broods was not significantly sex-biased (Fig. 1). Thus, flea infestations resulted in a significant bias in the proportion of male fledglings recruited, which could result from sex-related

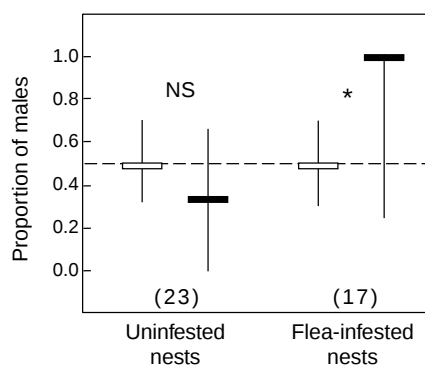


Figure 1 Flea infestation and proportion of males in great tit broods at fledging (white bars) and recruitment (black bars). Median values and interquartile ranges are shown. Number of broods recruiting young are in brackets. In infested broods, there was a significant difference in the proportion of males between fledging and recruitment (paired t -test: $t = -2.2$, d.f. = 16, $P = 0.04$; indicated by *). In uninfested broods the difference was not significant ($t = 1.8$, d.f. = 22, $P = 0.09$; indicated by NS). The change in the proportion of males between fledging and recruitment was significantly different between treatments (t -test: $t = 2.81$, d.f. = 38, $P = 0.007$).

differences in dispersal and/or mortality of fledglings after leaving the nest^{7,17,20}.

We analysed the data on all the recruits from the three years of experimental infestation to determine whether enhanced nestling competition within broods associated with flea infestation²¹ and the number of young in the brood^{22,23} affected the proportion of male recruits. We found a significant interaction between brood size at hatching and flea treatment on the proportion of male recruits in a brood (Fig. 2). In infested nests the proportion of male recruits tended to increase with brood size, whereas in uninfested broods the proportion of male recruits tended to decrease with brood size (Fig. 2). Our results indicate that increasing competition among nestlings in larger and infested broods lead to higher male local recruitment. Another study on great tits showed that an experimental enlargement of brood size also led to male-biased local recruitment²². In our study, variation in the effect of fleas with brood size on the proportion of male recruits could explain the fact that there was no overall significant effect of fleas on the total number of male and female recruits (uninfested, 29 males and 34 females; infested, 29 males and 19 females; $\chi^2_1 = 2.27$, $P = 0.13$).

We found no significant effect of fleas on the proportion of male fledglings when we analysed all broods sexed with DNA-markers ($Z = -0.53$, $P = 0.60$, $N = 37$ and 27), showing that fleas did not lead to sex-ratio biases in broods at fledging. In broods where we determined the sex of nestlings, males were heavier than females (mean difference at 14 days old: $0.7 \text{ g} \pm 0.1$; $N = 64$, $P < 0.001$). The sexual dimorphism in nestling body mass was not significantly affected either by fleas or by the number of young in the brood (fleas: $F(1, 61) = 0.32$, $P = 0.57$; brood size: $F(1, 61) = 0.03$, $P = 0.86$). Thus, differences in the proportion of males recruited were apparently not mediated by sex-dependent effects of fleas, or brood size on nestling body mass. There was no significant difference in the relative body mass of recruits in the two treatments (body mass deviation from the brood mean on day 14 (controlling for year and sex): $F(1, 106) = 0.77$, $P = 0.38$). This result indicates that young recruiting in the two treatments occupied similar relative positions in the body mass hierarchy within their broods²³.

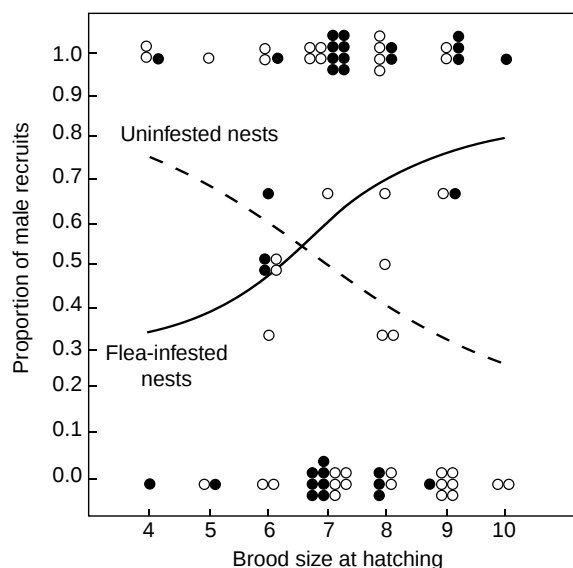


Figure 2 Interaction between flea infestation and brood size at hatching on the proportion of male recruits. Open symbols are for uninfested, and filled symbols for infested broods ($N = 75$). Fitted lines are from a logistic regression model on the number of male recruits (nominator) relative to the total number of recruits (denominator)²⁰, with year, flea treatment, brood size and their interactions as predictor variables. Only the interaction between flea treatment and brood size was significant (change in deviance of 5.49, d.f. = 1, $P = 0.02$; full model deviance was 101.8, d.f. = 71); all other effects were nonsignificant ($P > 0.20$).

We found that recruits from infested broods dispersed shorter natal distances (see Methods) than recruits from uninfested broods (mean for broods: infested, 532 ± 57 m; uninfested, 641 ± 44 m; generalized linear model: fleas, $F(1, 67) = 4.52$, $P = 0.04$; controlling for the effects of year, nestling mass and proportion of males (all $P > 0.10$). In contrast to other studies on great tits, which showed greater natal dispersal distances in females^{17,18,22}, the difference between the sexes was not significant in our study (mean for females, 660 ± 45 m, $N = 53$; mean for males, 547 ± 43 , $N = 58$; $F(1, 106) = 1.96$, $P = 0.16$; controlling for year and flea effects). The interaction between flea treatment and sex was not significant ($F(1, 105) = 0.27$, $P = 0.60$), indicating that fleas affected the natal dispersal distance of recruits independently of sex. Fleas could influence natal dispersal distances through an effect on nestling body mass^{18,22}, hormonal levels in nestlings²⁴ or through maternal effects^{8,12}. These different mechanisms suggest that fleas could have affected the competitive ability of individuals and their cost–benefit balance of dispersal^{3,25}. By dispersing shorter distances after growing in flea-infested nests, great tits remain closer to their natal territory where they could be better adapted or show higher tolerance to local parasites^{12,26,27}. This effect could be particularly important for females, as they face more frequent contacts with nest-based ectoparasites¹².

Current hypotheses for sex-biased dispersal propose that, in monogamous species with male resource defence like the great tit, males could gain greater benefits from their familiarity with essential local resources by remaining in their natal area^{3,10,28}. Our results of a male-biased recruitment in one year, and the interaction between flea treatment and brood size over three years, indicate that benefits of remaining in the natal area could be higher for males growing in large, flea-infested broods. A male-biased dispersal away from our study area from uninfested broods could also potentially explain lower local recruitment of males (Fig. 1). However, such male-biased dispersal would contrast with the female-biased dispersal pattern found in most monogamous birds^{3,10,17,28}.

To our knowledge, our study shows for the first time that parasite infestations affect host natal dispersal distances and can modify the sex-related probability of fledgling recruitment. Conditional sex-biased dispersal is not predicted by current theoretical models that focus mainly on inbreeding avoidance or local resource competition^{2,3,10,25,28}. Key concepts in sex-ratio theory and host–parasite coevolution have been developed by considering the effects of host dispersal^{1,4,26,27}. Our results indicate that, in certain organisms, hosts could modify their sex-allocation decisions in response to parasite infestations when parasites modify host sex-biased recruitment and dispersal. □

Methods

Experimental protocol. The study was performed in a population of great tits breeding in a mixed forest near Bern, Switzerland. Breeding pairs were assigned randomly to experimental treatments in each of three breeding seasons (1994–1996). In all three years, parasites were killed in all nests on the day the birds laid their second egg^{11–13}. Half the nests in the population were infested by a fixed number of adult fleas from the laying of the second egg or from the start of incubation onwards, whereas the other half were left uninfested. Uninfested nests in 1994 and 1995 were kept free of fleas by heat treatment in a microwave appliance every ten days; in 1996 nests were heat treated only once. Laying date and clutch size did not differ significantly between birds with flea-infested or uninfested nests (Wilcoxon two-sample test: laying date, $Z = -0.09$, $P = 0.93$; clutch size, $Z = 0.38$, $P = 0.71$; $N = 175$ and 168).

Data collection. For each breeding pair we recorded the date when the first egg was laid, clutch size, brood size at hatching, and the number of young fledged. Nestling body mass was measured to the nearest 0.1 g with an electronic balance on days 14 and 16 after hatching. On day 9, all nestlings were ringed with numbered aluminium rings. A local recruit was defined as a fledgling that was recaptured as a first-time breeder in the population until spring 1998. The linear distance between the nest where a young fledged to the nest where it was first re-captured as a breeder was taken as the natal dispersal distance¹⁰. In 1995,

nestlings of 64 broods were sexed using random amplified polymorphic DNA (RAPD) markers applied to the DNA extracted from blood collected in capillary tubes by venipuncture on 14-day-old nestlings²⁹. Mean values are shown $\pm$ standard errors. Analysis was performed in JMP and GLMStat.

Received 29 March; accepted 18 May 1999.

- Hamilton, W. D. Extraordinary sex ratios. *Science* **156**, 477–488 (1967).
- Hamilton, W. D. & May, R. M. Dispersal in stable habitats. *Nature* **269**, 578–581 (1977).
- Johnson, M. L. & Gaines, M. S. Evolution of dispersal: theoretical and empirical tests using birds and mammals. *Annu. Rev. Ecol. Syst.* **21**, 449–480 (1990).
- Herre, E. A. Population structure and the evolution of virulence in nematode parasites of fig wasps. *Science* **259**, 1442–1445 (1993).
- Clayton, D. H. & Moore, J. *Host–parasite Evolution: General Principles & Avian Models* (Oxford Univ. Press, Oxford, 1997).
- Richner, H. Host–parasite interactions and life-history evolution. *Zoology* **101**, 333–344 (1998).
- Brown, C. R. & Brown, M. B. Ectoparasitism as a cause of natal dispersal in cliff swallows. *Ecology* **73**, 1718–1723 (1992).
- Sorci, G., Massot, M. & Clobert, J. Maternal parasite load increases sprint speed and philopatry in female offspring of the common lizard. *Am. Nat.* **144**, 153–164 (1994).
- Danchin, E., Boulinier, T. & Massot, M. Conspecific reproductive success and breeding habitat selection: implications for the study of coloniality. *Ecology* **79**, 2415–2428 (1998).
- Greenwood, P. J. Mating systems, philopatry and dispersal in birds and mammals. *Anim. Behav.* **28**, 1140–1162 (1980).
- Heeb, P., Werner, I., Richner, H. & Kölliker, M. Horizontal transmission and reproductive rates of hen fleas in great tit nests. *J. Anim. Ecol.* **65**, 474–484 (1996).
- Heeb, P., Werner, I., Kölliker, M. & Richner, H. Benefits of induced host responses against an ectoparasite. *Proc. R. Soc. Lond. B* **265**, 51–56 (1998).
- Heeb, P., Kölliker, M. & Richner, H. Bird–ectoparasite interactions, nest humidity and ectoparasite communities. *Ecology* (in the press).
- Richner, H., Oppliger, A. & Christe, P. Effect of an ectoparasite on reproduction in great tits. *J. Anim. Ecol.* **62**, 703–710 (1993).
- Tinbergen, J. M. & Boerlijst, M. C. Nestling weight and survival in individual great tits (*Parus major*). *J. Anim. Ecol.* **59**, 1113–1128 (1990).
- Gebhardt-Heinrich, S. G. & Noordwijk, A. J. V. Nestling growth in the Great tit I. Heritability estimates under different environmental conditions. *J. Evol. Biol.* **3**, 341–362 (1991).
- Dhondt, A. A. & Hublé, J. Fledging-date and sex in relation to dispersal in young great tits. *Bird study* **15**, 127–134 (1968).
- Verhulst, S., Perrins, C. M. & Riddington, R. Natal dispersal of great tits in a patchy environment. *Ecology* **78**, 864–872 (1997).
- Poulin, R. Sexual inequalities in helminth infections: a cost of being a male? *Am. Nat.* **147**, 287–295 (1996).
- Brown, C. R., Brown, M. B. & Rannala, B. Ectoparasites reduce long-term survival of their avian host. *Proc. R. Soc. Lond. B* **262**, 313–319 (1995).
- Christe, P., Richner, H. & Oppliger, A. Begging, food provisioning, and nestling competition in great tit brood infested with ectoparasites. *Behav. Ecol.* **7**, 127–131 (1996).
- Smith, H. G., Källander, H. & Nilsson, J.-Å. The trade-off between offspring number and quality in the great tit *Parus major*. *J. Anim. Ecol.* **58**, 383–402 (1989).
- Kacelnick, A., Cotton, P. A., Stirling, L. & Wright, J. Food allocation among nestling starlings: sibling competition and the scope of parental choice. *Proc. R. Soc. Lond. B* **259**, 259–263 (1995).
- Belthoff, J. R. & Duffy, A. M. J. Corticosterone, body condition and locomotor activity: a model for natal dispersal in birds. *Anim. Behav.* **55**, 405–415 (1998).
- Weatherhead, P. J. & Forbes, M. R. L. Natal philopatry in passerine birds: genetic or ecological influences? *Behav. Ecol.* **5**, 426–433 (1994).
- Gandon, S., Capowiez, Y., Dubois, Y., Michalakakis, Y. & Olivieri, I. Local adaptation and gene-for-gene coevolution in a metapopulation model. *Proc. R. Soc. Lond. B* **263**, 1003–1009 (1996).
- Kaltz, O. & Shykoff, J. A. Local adaptation in host–parasite systems. *Heredity* **81**, 361–370 (1998).
- Clarke, A. L., Saether, B.-E. & Roskaft, E. Sex biases in avian dispersal: a reappraisal. *Oikos* **79**, 429–438 (1997).
- Lessells, C. M. & Mateman, A. C. Sexing birds using random amplified polymorphic DNA (RAPD) markers. *Mol. Ecol.* **7**, 187–195 (1998).
- Crawley, M. J. *GLIM for Ecologists* (Blackwell Science, Oxford, 1993).

Acknowledgements. We thank M. Lambrechts for a key suggestion, and T. C. M. Bakker, A. A. Dhondt, S. Heeb and A. Roulin for comments on the manuscript. This work was funded by grants to H.R. from the Swiss National Science Foundation.

Correspondence and requests for materials should be addressed to P.H. (e-mail: philipp.heeb@esh.unibe.ch).

Motion streaks provide a spatial code for motion direction

Wilson S. Geisler

Department of Psychology, University of Texas at Austin, Austin, Texas 78712, USA

Although many neurons in the primary visual cortex (V1) of primates are direction selective¹, they provide ambiguous information about the direction of motion of a stimulus^{2,3}. There is evidence that one of the ways in which the visual system resolves this ambiguity is by computing, from the responses of V1 neurons, velocity components in two or more spatial orientations

and then combining these velocity components²⁻⁹. Here I consider another potential neural mechanism for determining motion direction. When a localized image feature moves fast enough, it should become smeared in space owing to temporal integration in the visual system, creating a spatial signal—a 'motion streak'—oriented in the direction of the motion. The orientation masking and adaptation experiments reported here show that these spatial signals for motion direction exist in the human visual system for feature speeds above about 1 feature width per 100 ms. Computer simulations show that this psychophysical finding is consistent with the known response properties of V1 neurons, and that these spatial signals, when appropriately processed, are sufficient to determine motion direction in natural images.

Neurons in primate area V1 are selective for spatial orientation^{1,10}. Thus, a motion streak, even a short one, might produce a maximal response in the population of neurons whose preferred spatial orientation is parallel to the direction of motion. Such a relative maximum response across preferred orientation would be a robust spatial signal for motion direction. On the other hand, most cortical neurons are selective for motion perpendicular to their preferred spatial orientation^{1,10}; thus, a maximal response might occur in the population of neurons whose spatial orientation is approximately perpendicular to the direction of motion, at least for some feature speeds.

To determine whether spatial orientation signals for motion direction exist in the human visual system, luminance detection thresholds were measured for moving gaussian dots, as a function of dot size and speed, with dynamic random line masks oriented either parallel or perpendicular to the direction of dot motion. If a moving dot produces larger responses in the population of neurons whose orientation selectivity is parallel to the direction of motion then the parallel mask should be more effective in elevating threshold, because the relevant orientation-selective neurons would be very strongly activated by the mask. The two stimulus conditions are shown in Fig. 1. The target dot moved along the diagonal from the lower left towards the upper right, as indicated by the dashed arrow. A different random sample of line noise was presented in each frame, making the stimulus information for performing the task identical for the two mask orientations. Thresholds were measured for two subjects in a two-interval forced choice task.

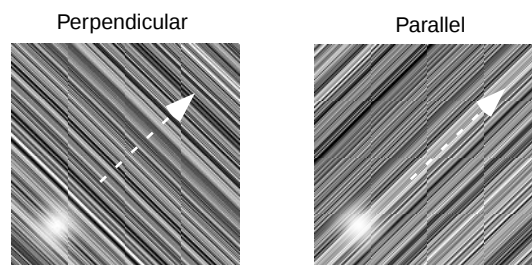


Figure 1 The two stimulus conditions in a masking experiment designed to determine whether the neural mechanisms most responsive to moving dots have their preferred spatial orientations parallel or perpendicular to the direction of motion. Each stimulus presentation consisted of a 25-frame (250-ms) movie. The target dot moved along the diagonal at a given speed, and was added to dynamic random line noise that was oriented either perpendicular or parallel to the direction of dot motion. The luminance threshold for dot detection was measured.

The data for one of the subjects are shown in Fig. 2a-d. Similar results were obtained for the other subject. Each panel shows the data for a different dot size. Within each panel, the vertical axis shows the luminance threshold (cd m^{-2}) for the moving dot, and the horizontal axis shows the speed of the dot motion in degrees per second. (Note that the scale changes from panel to panel.) When the dot was stationary (0° s^{-1}) the thresholds were identical, as expected given the circular symmetry of the dot. This result indicates that the differences in the effectiveness of the two mask orientations are not due to astigmatism or other optical factors. As the speed of movement of the dot increased, the thresholds remained the same until a critical dot speed was reached, beyond which the parallel mask became more effective. To estimate this critical speed, a four-parameter descriptive function was fitted to the data (solid curves in Fig. 2a-d). The descriptive function was specified by coordinates defining the right end-point of the horizontal line segment, and by a pair of slope parameters for the two rising segments of the data. Figure 2e shows the estimated critical speed (the end-point of the horizontal line segment) for the two subjects.

Apparently, whenever its speed exceeds about 1 dot width per 100 ms (dashed line), a moving dot produces a greater response in the population of neurons whose receptive fields are oriented

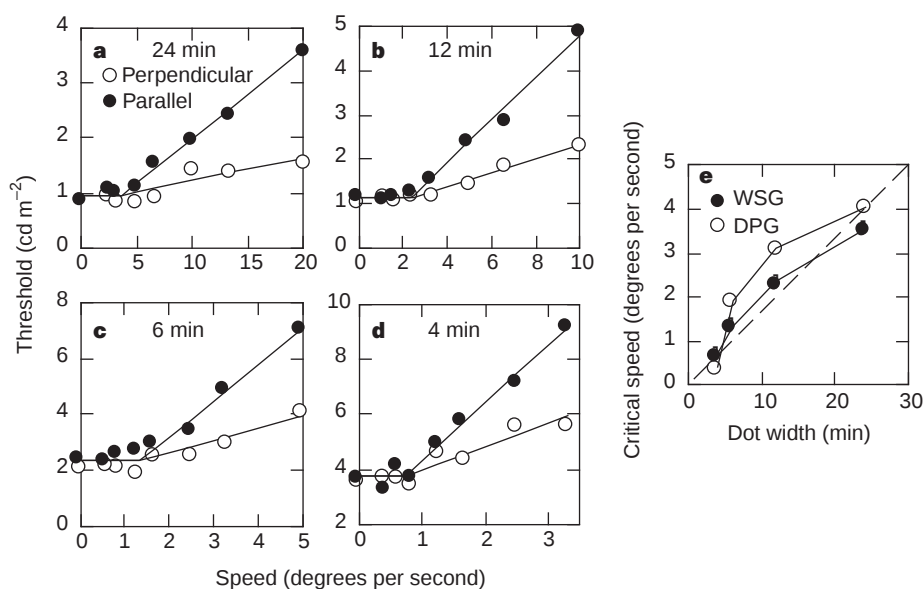


Figure 2 Results of forced-choice detection experiment. **a-d**, Luminance detection thresholds for target dots of 4, 6, 12 and 24-arcmin diameter, as a function of dot speed, for subject W.S.G. Solid symbols: line mask parallel to the direction of motion; open symbols: line mask perpendicular to the direction of

motion. Below a critical speed there is no effect of mask orientation; above that speed the parallel mask is more effective. **e**, Critical speed as function of dot width for the two subjects. Dashed line indicates 1 dot width per 100 ms.

parallel to the direction of motion. There seems to be no speed at which a greater response occurs in the population of neurons whose receptive fields are oriented perpendicular to the direction of motion. A second experiment, measuring detection threshold as a function of the orientation difference between the mask and the moving dot, showed that orientation differences between 0° (parallel) and 90° (perpendicular) produce intermediate levels of threshold elevation. Thus, the experimental results indicate that there is a robust and stable spatial orientation signal for motion direction in the visual system for a wide range of speeds, down to less than 1° s^{-1} for the smallest dot size.

I used computer simulations to determine whether these psychophysical results are consistent with what would be expected from the responses of V1 neurons. The responses of cortical 'simple cells' and 'complex cells' were simulated using the mean values of the tuning characteristics for V1 neurons reported elsewhere (see Methods)¹¹. The preferred spatial frequency of the simulated neurons was set to the value that produced the greatest response to a 6-min dot (9 cycles per degree). Figure 3a shows, as a function of the speed of the moving dot, the peak response for motion parallel to the preferred spatial orientation divided by the peak response for motion perpendicular to the preferred spatial orientation. The solid circles show this response ratio for a dot trajectory passing through the centre of the receptive field of a non-direction-selective cell. Note that the psychophysically measured critical speed for this dot size (indicated by the vertical arrow) is near the speed at which the response ratio begins to exceed 1.0. The triangles show the

response ratio for a direction-selective cell whose preferred direction corresponds to the direction of dot motion. For this cell the response ratio is much less than 1.0 at low speeds. However, the response ratio of the summed responses of two direction-selective neurons with opposite motion-direction preferences is similar to the response ratio of the non-direction-selective cell. Furthermore, essentially identical results were obtained for simple and complex cells. Thus, the simulations indicate that the psychophysical results are consistent with the population response expected to be obtained by pooling across V1 neurons with the same preferred spatial orientation.

Although the masking experiments indicate that a spatial orientation signal for motion direction exists in the human visual system, two important concerns remain. The first is that motion streaks provide ambiguous information about the direction of motion. A detected streak could be the result of a moving feature, or it could be a static image contour. Even if it is the result of a moving feature, the direction in which the feature moved along the orientation of the streak is uncertain. The second concern is whether the spatial orientation signals would actually be very useful for determining motion direction in natural images, because natural images generally contain an extremely complex mixture of extended contours and localized features.

These concerns were addressed by considering the performance of the spatial motion-direction sensor shown in Fig. 3b. The proposed sensor combines multiplicatively an oriented 'non-direction-selective cell' with a perpendicularly oriented 'direction-selective (motion

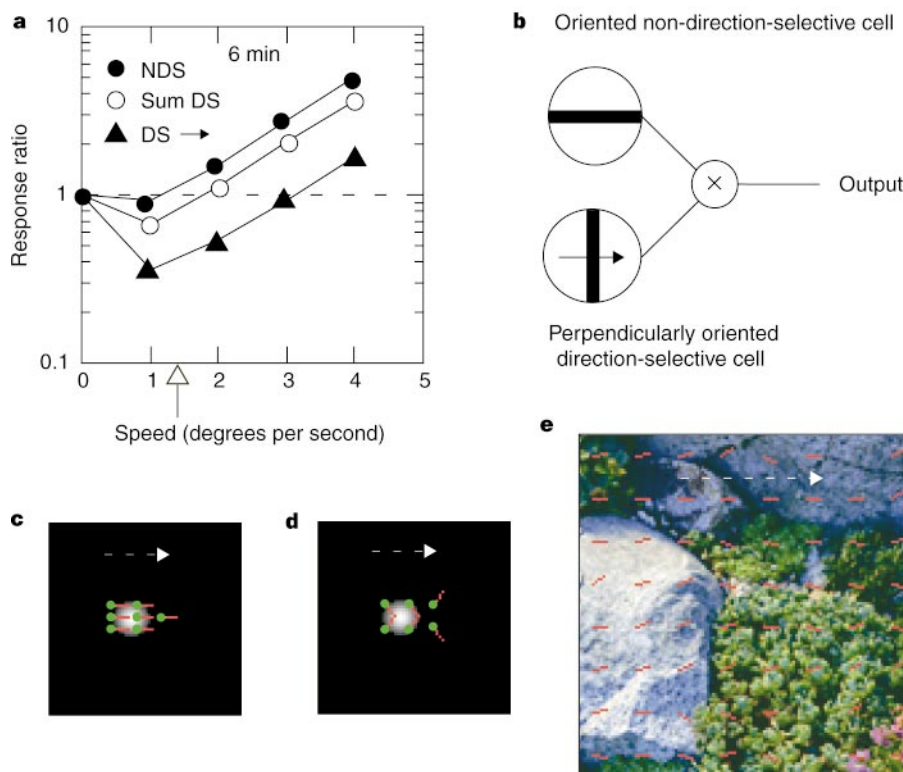


Figure 3 Neural simulations. **a**, Computed responses of typical cortical neurons in primate V1 for a 6-min dot moving parallel or perpendicular to the preferred spatial orientation of the neuron. The ratio of the response in the parallel direction to that in the perpendicular direction is plotted as a function of dot speed. Results are shown for a non-direction-selective cell (NDS), a cell that prefers motion to the right (DS $\rightarrow$) and the sum of the responses of a cell that prefers motion to the right with one that prefers motion to the left (sum DS). **b**, A possible neural sensor for determining motion direction from motion streaks. An oriented non-direction-selective cell (or population of cells) is combined multiplicatively with a perpendicularly oriented direction-selective cell (or population of cells). **c**, Responses of a field of spatial motion direction sensors (of all preferred

orientations) to a 6-min Gaussian dot moving at 4° s^{-1} in the direction of the arrow. For each location indicated by a green dot, the red line segment shows the orientation of the motion sensor which produced the largest response. **d**, Responses of a field of direction-selective cells alone. The line segments are drawn perpendicular to the orientation of the cells which produced the largest response. **e**, Responses of a field of spatial motion sensors to a $1^\circ \times 1^\circ$ natural image moving in the direction of the arrow. The orientations of the red line segments indicate the estimated direction of motion in each 7.5-min square block. The direction estimates are given by $\hat{\theta} = \arctan[\sum R_i \sin \theta_i / \sum R_i \cos \theta_i]$, where θ_i is the preferred spatial orientation of the i th sensor at the given location, and R_i is the response of the sensor.

opponent¹²) cell'. For this sensor, localized features moving to the right in the preferred orientation of the non-direction-selective cell will produce a large response, whereas stationary features or features moving to the left will produce no response. The simulations described below show that the sensor produces a more accurate and stable signal for the motion direction of localized features than does the typical direction-selective cell in V1.

Two versions of the sensor were tested. In the first version, the response of the non-direction-selective cell was directly multiplied by the response of the direction-selective cell. The second version was designed to evaluate the usefulness of the spatial motion signals by using as little information as possible from the direction-selective cells. In this case, the response of the direction-selective cell was applied as a simple logical operator: if the response was positive, indicating motion in one general direction (for example, 0–180°), the non-direction-selective cell response was multiplied by one; otherwise, it was multiplied by zero. The two versions gave similar results. The performance of the second version is shown here.

Figure 3c shows the response of a dense array of motion-direction sensors to a 6-min gaussian dot moving from left to right at 4° s⁻¹. The sensor array consisted of 36 sensors (one for every 10° orientation) centred on each image pixel. Each red line segment indicates the preferred orientation of the sensor that gave the largest response at the spatial location indicated by the green dot. The sensors that gave the largest responses corresponded to the true direction of motion. This behaviour holds over a wide range of dot velocities. For comparison, Fig. 3d shows the response of a similar array of direction-selective cells. At this dot speed, the cells that gave the largest response corresponded to either 60° or -60° from the true direction of motion. In general, as dot speed increases, the largest responses correspond to directions that are increasingly far from the true direction of motion.

Figure 3e shows the motion direction estimated at each location in a 1° × 1° natural image moving at 4° s⁻¹ to the right. The natural image movie was created by sliding (in effect) a larger image past a 128 × 128-pixel viewing aperture. Although the motion direction is not estimated accurately everywhere in the image, the estimates are generally good. Similar results were obtained for other natural images. Thus, it seems safe to conclude that there is almost always useful motion-direction information encoded by spatial mechanisms in the visual cortex once critical speed is exceeded.

The psychophysical results and the simulations show that moving image features produce robust spatial orientation signals in the human visual system, even at low speeds, and that these signals contain useful information about motion direction. Although it is unlikely that this information would be ignored by the brain, the results so far do not prove that motion streaks are used in determining motion direction. Some preliminary evidence that they are used for this purpose was obtained in an orientation adaptation experiment. Two subjects adjusted the apparent direction of a moving 12-min dot (on a dark background) until it appeared to move vertically, after adaptation to a 4 cycles per degree grating oriented either 10° to the left, 10° to the right, or vertically. The measurements for left and right adaption were combined and subtracted from the vertical baseline measurements. Figure 4a shows the results for dot speeds of 2.5° s⁻¹ and 10° s⁻¹, which should produce weak and strong motion streaks, respectively (Fig. 2b). If motion streaks are being used to determine motion direction then orientation adaptation should have a stronger effect on the apparent motion direction of the faster moving dot. This is what was found. An alternative explanation for the result might be based upon the fact that the slower-moving dot covered a smaller total distance. However, this was ruled out with measurements for static dots located at the end points of the motion path. As can be seen in Fig. 4b, the shift is approximately the same for both dot separations.

There is evidence that the visual system contains mechanisms for

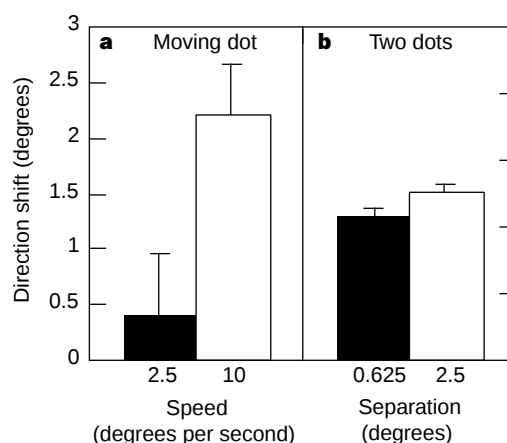


Figure 4 Results of tilt after-effect experiment. **a**, Effect of adaptation to a grating tilted 10° from vertical on the perceived direction of a 12-min dot moving vertically, at one of two speeds. **b**, Effect of adaptation on the perceived orientation of a static pair of dots oriented vertically, at one to two spatial separations. The vertical axis represents the shift in motion direction or orientation required to cancel the adaptation effect.

suppressing the apparent blur due to motion^{13–16}. Such mechanisms are thought to be valuable for preserving the apparent spatial structure of objects across different speeds (for example, maintaining shape constancy). By contrast, my data show that motion blur produces substantial responses in orientation-selective mechanisms, even at low speeds, and that these spatial signals are used for encoding moving objects. These two aspects of motion blur are not incompatible; the spatial motion signals in area V1 may be amplified in areas of the brain responsible for encoding motion, but suppressed in areas of the brain responsible for shape and object representation. This hypothesis is consistent with recent demonstrations that blur discrimination performance is uncorrelated with the reduced apparent blur of moving stimuli¹⁷.

In conclusion, the psychophysical experiments and physiological analyses reported here indicate that, in addition to estimating motion direction by combining velocity components, the human visual system may also estimate motion direction by performing a spatial analysis of motion streaks. An advantage of this scheme is that the two approaches for estimating motion direction are complementary: as speed increases, estimates of velocity components become less reliable, while the estimates of the orientation of motion streaks becomes more reliable. □

Methods

Experiment. The subjects were two experienced psychophysical observers with normal acuity. The stimuli were displayed on a Philips monitor (Model 201B) run at a frame rate of 100 Hz and a spatial resolution of 640 × 480, using only the green gun ($x = 0.295$, $y = 0.580$). At the viewing distance of 236 cm, the 480 × 480-pixel random mask was 6.8° × 6.8° in visual angle. Each frame of the random mask consisted of lines one pixel wide that were randomly assigned a luminance of either 0.8 cd m⁻² or 1.92 cd m⁻². For subject D.P.G., the random line masks were diagonal, as in Fig. 1; for subject W.S.G., who was run first, the random line masks were vertical and horizontal. (Diagonal masks were used for D.P.G. because pretesting revealed a slightly greater effect of vertical maskers when the dots were stationary.) The dot target was a radially symmetrical two-dimensional gaussian and was added to the random line mask. In the plots, the dot width is defined as four times the standard deviation of the gaussian.

Four-point psychometric functions were measured with a two-interval blocked forced choice procedure, using the method of constant stimuli. A different dynamic random mask appeared in each interval and the target dot was added to one of the two intervals at random. The durations of the mask and the dot were both 250 ms. The amplitude of the dot was ramped on in the first

50 ms and off in the last 50 ms by a raised cosine function. The two presentation intervals were separated by 500 ms, and the trials were separated by 2,300 ms. Feedback was provided on each trial. Five blocks of 30 trials were run for each psychometric function. To allow for adaptation to the random line masks and to provide some practice, the amplitude of the dot in the first and second blocks was made the same, and the responses from the first block were discarded. The data points in Fig. 2 represent the average of two thresholds estimated from psychometric functions measured a number of days apart. The order of conditions in the repeated session was opposite to that in the first session, to minimize fatigue/practice artefacts.

Simulations. The simulations involved computing the responses of model neurons to movies consisting of 32 frames of 128×128 pixel images, representing a movie duration of 0.25 s and an image area of $1^\circ \times 1^\circ$. Generally, there were 36 model neurons (one for every 10° orientation) centred on each image pixel location. In different simulations, the model neurons were 'simple cells', 'complex cells', 'motion opponent cells' or the proposed 'motion-direction sensors'. Most of the receptive field characteristics of the model neurons were fixed at the average values reported for the monkey¹¹: orientation bandwidth, 40° ; spatial frequency bandwidth, 1.5 octaves; preferred temporal frequency, 8 Hz; temporal frequency bandwidth, 2.5 octaves; nonlinear response exponent, 2. The preferred spatial frequency of the model neurons was kept at 9 cycles per degree, and the direction-selectivity index (1 – non-preferred/preferred) of the direction-selective cells at 0.9. (The bandwidths and direction selectivity refer to the effective values after applying the response exponent.)

The methods of computing responses were similar to those described in refs 5 and 12. The convolutions for each orientation were computed separately using Fourier methods. The component shapes of the spatial transfer function were a gaussian function on a log axis (the so-called log Gabor function) for one spatial direction, and a gaussian function centred on the origin for the other spatial direction. The component shape of the temporal transfer function was the transfer function of a difference-of-gammas function. (Component shapes corresponding to the Hilbert transforms of the log Gabor and difference-of-gammas functions were also used.) The degree of direction selectivity was controlled by combining these components with an appropriate weighting factor^{5,12}.

After multiplication of the transformed movie by the transfer function, the result was inverse transformed. To obtain 'simple cell' responses, the log Gabor component of the transfer function was set to cosine phase; the inverse transform was half-wave rectified and squared. To obtain 'complex cell' responses, the inverse transform was computed separately with the log Gabor component in cosine phase and in sine phase; the cosine-phase and sine-phase responses were then squared and summed^{5,12}. 'Motion opponent cell' responses were obtained by differencing the responses of complex cells having opposite preferred directions of motion and the same preferred spatial orientation^{5,12}.

Received 27 January; accepted 26 April 1999.

- Hubel, D. H. & Wiesel, T. N. Receptive fields and functional architecture of monkey striate cortex. *J. Physiol. (Lond.)* **195**, 215–243 (1968).
- Marr, D. & Ullman, S. Direction selectivity and its use in early visual processing. *Proc. R. Soc. Lond. B* **212**, 151–180 (1981).
- Adelson, E. H. & Movshon, J. A. Phenomenal coherence of moving visual patterns. *Nature* **300**, 523–525 (1982).
- Albright, T. D. Direction and orientation selectivity of neurons in visual area MT of the macaque. *J. Neurophysiol.* **52**, 1106–1130 (1984).
- Watson, A. B. & Ahumada, A. J. Model of human visual-motion sensing. *J. Opt. Soc. Am. A* **2**, 322–342 (1985).
- Simoncelli, E. P. & Heeger, D. J. A model of neuronal responses in visual area MT. *Vision Res.* **38**, 743–761 (1998).
- Smith, A. T. & Snowden, R. J. (eds) *Visual Detection of Motion* (Academic, London, 1994).
- Derrington, A. & Suero, M. Motion of complex patterns is computed from the perceived motions of their components. *Vision Res.* **31**, 139–149 (1991).
- Stone, L. S., Watson, A. B. & Mulligan, J. B. Effect of contrast on the perceived direction of a moving plaid. *Vision Res.* **30**, 1049–1067 (1990).
- De Valois, R. L., Yund, E. W. & Hepler, N. The orientation and direction selectivity of cells in macaque visual cortex. *Vision Res.* **22**, 531–544 (1982).
- Geisler, W. S. & Albrecht, D. G. Visual cortex neurons in monkeys and cats: Detection, discrimination, and identification. *Vis. Neurosci.* **14**, 897–919 (1997).
- Adelson, E. H. & Bergen, J. R. Spatiotemporal energy models for the perception of motion. *J. Opt. Soc. Am. A* **2**, 284–299 (1985).
- Anderson, C. H. & Van Essen, D. C. Blur into focus. *Nature* **343**, 419–420 (1990).
- Morgan, M. J. & Benton, S. Motion-deblurring in human vision. *Nature* **340**, 385–386 (1989).
- Ramachandran, V. S., Madhusudan Rao, V. & Vidyasagar, T. R. Sharpness constancy during movement perception. *Perception* **3**, 97–98 (1974).
- Burr, D. Motion smear. *Nature* **284**, 164–165 (1980).
- Burr, D. C. & Morgan, M. J. Motion deblurring in human vision. *Proc. R. Soc. Lond. B* **264**, 431–436 (1997).

Acknowledgements. I thank B. Henning for pointing out the potential value of orientation masking in this context. D. Albrecht, L. Cormack and B. Henning provided helpful discussions as well as comments on the manuscript. Supported by the National Eye Institute, NIH.

Correspondence and requests for materials should be addressed to W.S.G. (e-mail: geisler@psy.utexas.edu).

Mox2 is a component of the genetic hierarchy controlling limb muscle development

Baljinder S. Mankoo*, Nina S. Collins*, Peter Ashby†‡, Elena Grigorieva*, Larysa H. Pevny‡‡, Albert Candia§§, Christopher V. E. Wright§, Peter W. J. Rigby† & Vassilis Pachnis*

Divisions of *Developmental Neurobiology, †Eukaryotic Molecular Genetics and ‡Developmental Genetics, MRC National Institute for Medical Research, The Ridgeway, Mill Hill, London NW7 1AA, UK

§Vanderbilt Medical School, B2317A MCN, 1161 21st Avenue South, Nashville, Tennessee 37232-2175, USA

The skeletal muscles of the limbs develop from myogenic progenitors that originate in the paraxial mesoderm and migrate into the limb-bud mesenchyme¹. Among the genes known to be important for muscle development in mammalian embryos are those encoding the basic helix-loop-helix (bHLH) myogenic regulatory factors (MRFs; *MyoD*, *Myf5*, *myogenin* and *MRF4*)^{2–4} and *Pax3*, a paired-type homeobox gene that is critical for the development of limb musculature^{5–7}. *Mox1* and *Mox2* are closely related homeobox genes that are expressed in overlapping patterns in the paraxial mesoderm and its derivatives^{8,9}. Here we show that mice homozygous for a null mutation of *Mox2* have a developmental defect of the limb musculature, characterized by an overall reduction in muscle mass and elimination of specific muscles. *Mox2* is not needed for the migration of myogenic precursors into the limb bud, but it is essential for normal appendicular muscle formation and for the normal regulation of myogenic genes, as demonstrated by the downregulation of *Pax3* and *Myf5* but not *MyoD* in *Mox2*-deficient limb buds. Our findings show that the MOX2 homeoprotein is an important regulator of vertebrate limb myogenesis.

Mox1 and *Mox2* are co-expressed in the epithelial somites of mouse embryos^{8,9} (Fig. 1a, c). In differentiating somites, both genes continue to be expressed in the sclerotome, but the dermomyotome expresses only *Mox1*^{8,9} (Fig. 1b, d, i). Limb buds also show differential regulation of *Mox* genes, in which *Mox2* (but not *Mox1*) messenger RNA was detected in a pattern consistent with expression in the migrating myoblasts and their derivative muscles⁹ (Fig. 1e–h). This was further demonstrated by double labelling with immunohistochemistry for MOX2 and *in situ* hybridization histochemistry for *Pax3*, a gene that is expressed in limb buds, specifically in the dermomyotome-derived myoblasts^{5–7,10}. *Pax3* and *Mox2* were co-expressed in most migrating myoblasts as soon as they de-epithelialized from the dermomyotome and entered the limb-bud mesenchyme (Fig. 1i–j). To identify further the *Mox2*-expressing cells in limb buds, we analysed the expression of *Mox2* in homozygous *Splotch* (*Sp*) embryos, in which a mutation at the *Pax3* locus prevents the migration of myogenic precursors from their site of origin in the paraxial mesoderm into the limb buds^{5,6,10}. *Mox2* transcripts, which are seen in scattered cells in the proximal region of forelimbs and hindlimbs of embryonic day 10.5 (E10.5) $+/+$ or $+/-Sp$ embryos, were absent from the limbs of their *Sp/Sp* littermates (Fig. 1k–l and data not shown). These data show that, in the

‡Present addresses: Wellcome Trust Building, University of Dundee, WTB/MSI Complex, Dow Street, Dundee DD1 5EH, UK (P.A.); Developmental Genetics Programme, Krebs Institute, University of Sheffield, Sheffield S10 2NT, UK (L.H.P.); Department of Developmental Biology, Stanford University, Stanford, California 94305, USA (A.C.).

myogenic progenitors of limb buds, *Mox1* and *Mox2* have reciprocal expression patterns. *Mox1* (but not *Mox2*) is expressed in the dermomyotome where the limb muscle progenitors originate, but after de-epithelialization and invasion of the limb-bud mesenchyme they express *Mox2* (but not *Mox1*). Expression of *Mox2* was also detected in a distal limb-bud domain that, being unaffected by the *Sp* mutation, is likely to represent mesenchyme derived from lateral plate mesoderm (Fig. 1l).

To evaluate the role of *Mox2* in the development of paraxial mesoderm and its derivatives, we disrupted its first coding exon by gene targeting in mouse embryonic stem (ES) cells to generate a null allele (Fig. 2). Mice heterozygous for this mutation (*Mox2*^{+/-}) displayed no abnormalities on a hybrid C57B1/6-129/Ola background and homozygous mutant animals (*Mox2*^{-/-}) were born with the expected mendelian frequency, indicating no embryonic lethality. However, newborn *Mox2*^{-/-} animals exhibited characteristic limb defects, manifested as an abnormal extended position of the forelimbs and hindlimbs and markedly reduced motility (Fig. 3a). This phenotype was variable; most animals (80%) were severely affected, failed to thrive and died before weaning. Some of the least affected animals survived to adulthood, displayed abnormal gait and showed a reflex which is frequently observed in mice with muscular dystrophy¹¹ and consists of joint contractures and hind-leg grasping upon lifting by the tail (Fig. 3b, c). In addition to the limb defects, about 10% of *Mox2*^{-/-} animals had clefting of the secondary palate, consistent with *Mox2* expression in this tissue (data not shown).

Histological sections showed that *Mox2*^{-/-} neonates had an

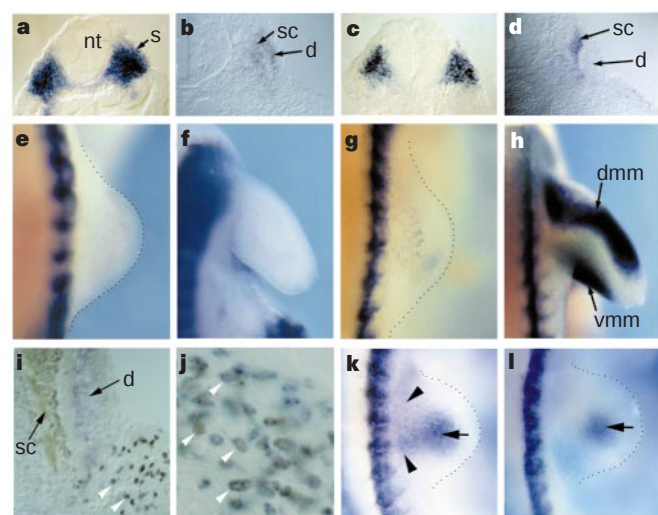


Figure 1 Expression of *Mox1* and *Mox2* in the paraxial mesoderm and the limb buds of mouse embryos. **a-d**, Transverse sections of whole-mount preparations of E9.5 embryos hybridized with riboprobes specific for *Mox1* (**a**, **b**) and *Mox2* (**c**, **d**). **a**, **c**, Caudal epithelial somites; **b**, **d**, are from the forelimb bud level. Newly formed epithelial somites (s), dermomyotome (d), sclerotome (sc) and neural tube (nt) are shown. **e-h**, Expression of *Mox1* (**e**, **f**) and *Mox2* (**g**, **h**) in limb buds of E9.75 (**e**, **g**; dorsal views) and E10.75 (**f**, **h**; posterior views) embryos. *Mox2* mRNA is detected in the somite-derived myoblasts that populate the limb bud and coalesce to form the dorsal and ventral muscle masses (dmm, vmm). *Mox1* mRNA is not detected in the limb bud at these stages (**e**, **f**). Transverse sections of E10.5 embryos double-labelled for *MOX2* and *Pax3* (**i**, **j**) show non-overlapping expression of *MOX2* protein (brown) in the sclerotome and *Pax3* mRNA (purple) in the dermomyotome. Myoblasts migrating from the dermomyotome into the limb bud (**i**, arrowheads) and most migrating myoblasts in the limb bud (**j**, arrowheads) express both *Mox2* and *Pax3*. **k**, **l**, Expression of *Mox2* mRNA in forelimb buds of E10.5 control (**k**) and *Splootch*^{2H} homozygous (**l**) embryos. There is proximal expression in control forelimbs (**k**, arrowheads) but not in mutants (**l**). A more distal domain of expression (arrows) of *Mox2* was present in both control and mutant limb buds.

overall reduction in skeletal muscle mass of the limbs that was associated with the absence of specific muscles (Fig. 3d-i). For example, in the proximal forelimb, *musculus (m.) deltoideus*, the lateral head of *m. triceps brachii*, *m. teres major* and *m. biceps brachii* were absent or significantly reduced in size, whereas the long and medial heads of *m. triceps brachii* were unaffected (Fig. 3d, e). In the distal forelimb, several muscles of the flexor compartment, such as *m. flexor digitorum profundus*, *m. flexor carpi ulnaris* and *m. flexor carpi radialis* were absent (Fig. 3f, g). The extensor muscle group showed similar defects. Although no muscles were missing in the hindlimbs, there was a marked decrease in the overall muscle mass, and in particular that of *m. gastrocnemius* (Fig. 3h, i). In addition, there were many myopathic skeletal muscle fibres, characterized by variable diameter and centrally placed nuclei, in both fore- and hindlimbs (Fig. 3j, k). No major anatomical or histological defects were detected in the body-wall muscles or in the tongue muscles, which are derived from the hypoglossal cord¹² (data not shown). Finally, the axial and appendicular skeleton appeared normal (data not shown), indicating that the muscle defects were not secondary to skeletal abnormalities.

To determine whether the muscle abnormalities of *Mox2*^{-/-} neonates result from a developmental defect, β -galactosidase (β -gal) histochemistry was used to examine *Mox2*^{-/-} embryos carrying a myogenin-nLacZ reporter transgene (GNZ). GNZ, which contains a gene encoding nuclearly localized β -gal under the control of the *myogenin* promoter, mimics the expression of the endogenous *myogenin* locus and is transcribed in all postmitotic (and, in the limb, postmigratory) skeletal muscle fibres¹³⁻¹⁵. At the early stages of limb muscle differentiation (E11.5-12.5), despite normal expression of GNZ in the trunk and head muscles (Fig. 4a), the number of β -gal⁺ cells in the fore- and hindlimbs of *Mox2*^{-/-} embryos was significantly reduced (Fig. 4b, c). At later stages, when the identification of individual muscles was feasible, we observed, in addition to the reduced number of GNZ-expressing cells, alterations in the pattern

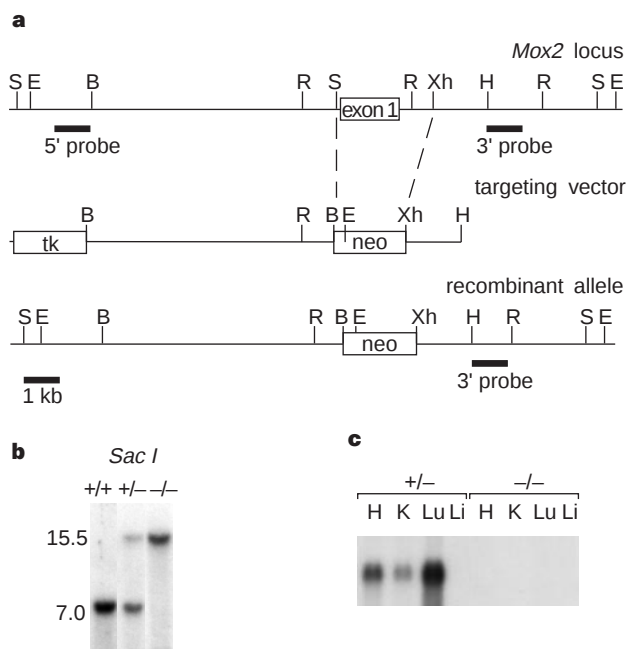


Figure 2 Targeted disruption of the *Mox2* locus. **a**, Partial restriction maps of the wild-type *Mox2* locus (top), targeting vector (middle) and recombinant allele (bottom). B, *Bam*HI; E, *Eco*RV; H, *Hind*III, R, *Eco*RI; S, *Sac*I; Xh, *Xho*I. **b**, Southern blot analysis of *Sac*I-digested genomic DNA isolated from wild-type (+/+), heterozygous (+/-) and homozygous mutant (-/-) mice. The Southern blot was hybridized with a 3' flanking external probe. **c**, Northern blot analysis of RNA isolated from organs of heterozygous (+/-) and homozygous mutant (-/-) neonates using a *Mox2*-specific riboprobe derived from the 3'UTR sequences. H, heart; K, kidney; Lu, lung; Li, liver.

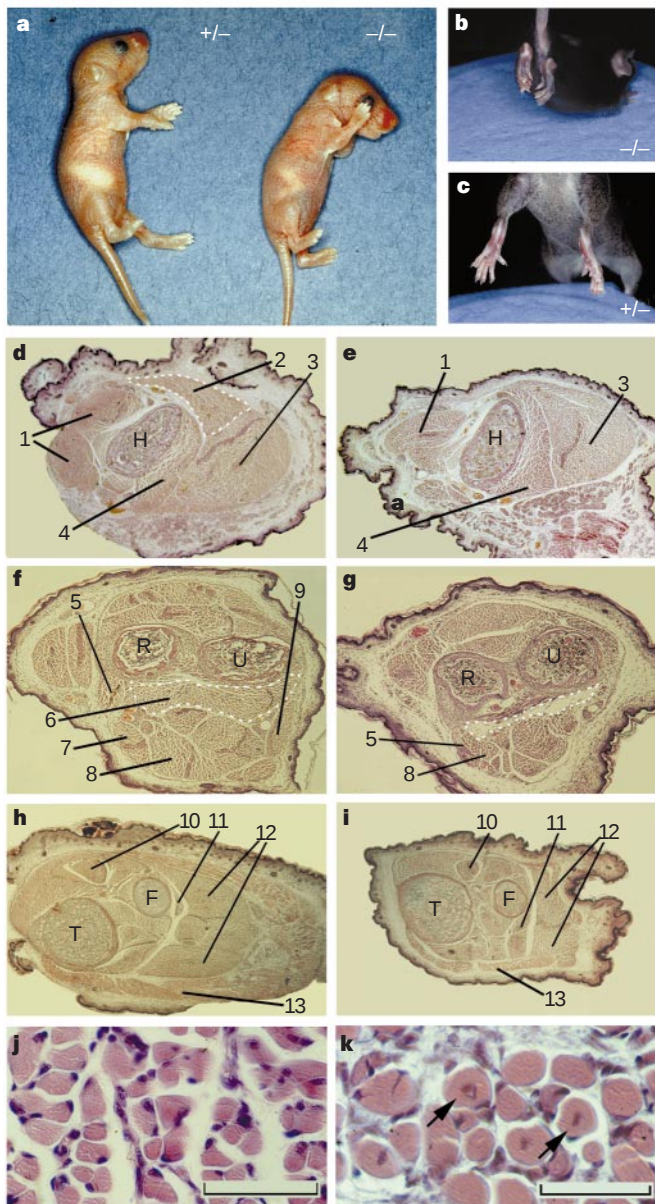


Figure 3 Limb muscle defects of *Mox2*^{-/-} homozygous mutants. **a**, Homozygous mutant (-/-, *n* = 10) neonates display abnormal position of forelimbs and hind limbs compared with control (+/-, *n* = 10) littermates. An abnormal grasping reflex observed in mutant animals (**b**) was absent from control littermates (**c**). **d-k**, Histological analysis of limbs from control (**d, f, h, j**, *n* = 2) and mutant *Mox2*^{-/-} (**e, g, i, k**, *n* = 4) neonates. Transverse paraffin sections were stained with haematoxylin and eosin. **d, e**, Sections through the proximal forelimb; **f, g**, Sections through the distal forelimb; **h, i**, Sections distal to the knee of the hindlimb. H, humerus; R, radius; U, ulna; T, tibia; F, fibula. Individual muscles indicated by numbers: 1, *m. biceps brachii*; 2, *m. triceps brachii* (lateral head); 3, *m. triceps brachii* (long head); 4, *m. triceps brachii* (medial head); 5, *m. pronator teres*; 6, *m. flexor digitorum profundus*; 7, *m. flexor carpi radialis*; 8, *m. flexor digitorum superficialis*; 9, *m. flexor carpi ulnaris*; 10, *m. extensor digitorum longus*; 11, *m. soleus*; 12, *m. gastrocnemius*; 13, *m. biceps femoris*. Outlined areas in **d, f** and **g** indicate the position of the missing *m. triceps brachii* (lateral head; **d**) and *m. flexor digitorum profundus* (**f, g**). Muscle anatomy was confirmed by dissection of adult limbs, serial sectioning of neonate limbs and analysis of *GNZ/Mox2* embryos at E13.5–15.5 **j, k** High-power photomicrographs of sections of a region of hindlimb muscle from control (**j**) and *Mox2*^{-/-} (**k**) neonates. Note the centrally placed nuclei (arrows) and variable diameter fibres in mutant muscles. Scale bar, 200 μ m.

of limb muscle formation which prefigured the defects observed in the limbs of newborn animals. For example, in E13.5 *GNZ/Mox2*^{-/-} embryos, *m. deltoideus*, *m. teres major* and the lateral head of *m. triceps brachii* were absent from the forelimbs (Fig. 4d, e). At this stage, we consistently observed abnormal splitting of the *m. extensor digitorum longus* of the hindlimbs (Fig. 4f, g). These findings indicate that the muscle defects in the limbs of newborn *Mox2*^{-/-} animals can be traced back to early stages of muscle formation, and that the MOX2 homeoprotein is critical for normal morphogenesis of limb musculature.

To analyse the developmental defects of the limb musculature in *Mox2*^{-/-} mice further, we examined the expression of several genes that function during various stages of limb myogenesis. First, we studied the expression of *Lbx1*, *c-Met* and *Pax3*, three genes expressed in limb myoblasts from their origin at the lateral dermo-myotome and throughout their migration within the limb-bud mesenchyme^{5–7,16,17}. Comparable levels of expression of all three genes were observed in the trunk (dermo-myotome and neural tube) of control and mutant E10.5 embryos (Fig. 5). Similar levels and patterns of expression of *Lbx1* were also observed in limb buds irrespective of genotype (Fig. 5a, b), whereas *c-Met* expression was slightly reduced in mutant limb buds (Fig. 5c, d). In contrast, there was a marked downregulation of *Pax3* in limb buds of E10.5 *Mox2*^{-/-} embryos relative to control littermates (Fig. 5e, f). Analysis of younger (E9.75–10.0) mutant embryos (Fig. 5g, h) revealed that there was reduced expression of *Pax3* as myoblasts delaminated from the epithelial dermo-myotome, which was downregulated during further migration. We next examined the expression of myogenic bHLH genes, which are expressed after migration. In

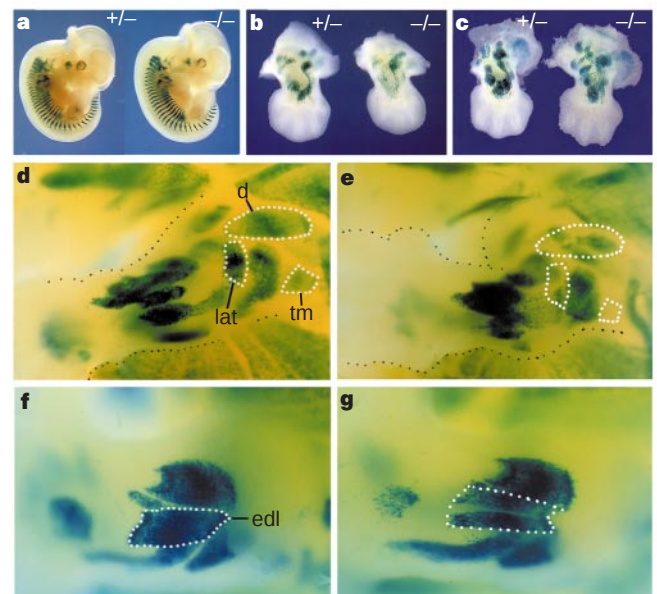


Figure 4 Limb muscle morphogenesis is disrupted in *Mox2*^{-/-} embryos. A myogenin-nLacZ transgene (*GNZ*) was used as a reporter to visualize individual muscles during embryogenesis. **a**, β -galactosidase histochemistry on whole mount preparations of E11.5 control (+/-, *n* = 4) and mutant (-/-, *n* = 4) embryos shows no defects in body wall musculature. In contrast, there was a marked and reproducible reduction in the overall number of β -gal⁺ fibres in mutant (-/-) forelimbs at E11.5 (**b**) and hindlimbs at E12.5 (**c**), compared with controls (+/-). Only dorsal muscle masses are shown in **b** and **c**; we obtained similar results for ventral limb-bud muscles (not shown). **d-g**, Dorsal aspects of left forelimbs (**d, e**) and hindlimbs (**f, g**) of control (**d, f**, *n* = 4) and mutant (**e, g**, *n* = 4) E13.5 embryos. The muscles ablated in mutant forelimbs are outlined and labelled; d, *m. deltoideus*; lat, *m. triceps brachii* (lateral head); tm, *m. teres major*. In the absence of the outlined *m. deltoideus*, the underlying muscles are revealed (**e**). Abnormal splitting of *m. extensor digitorum longus* (edl) (outlined in **f** and **g**) was consistently observed in mutant (**g**) relative to control (**f**) hindlimbs.

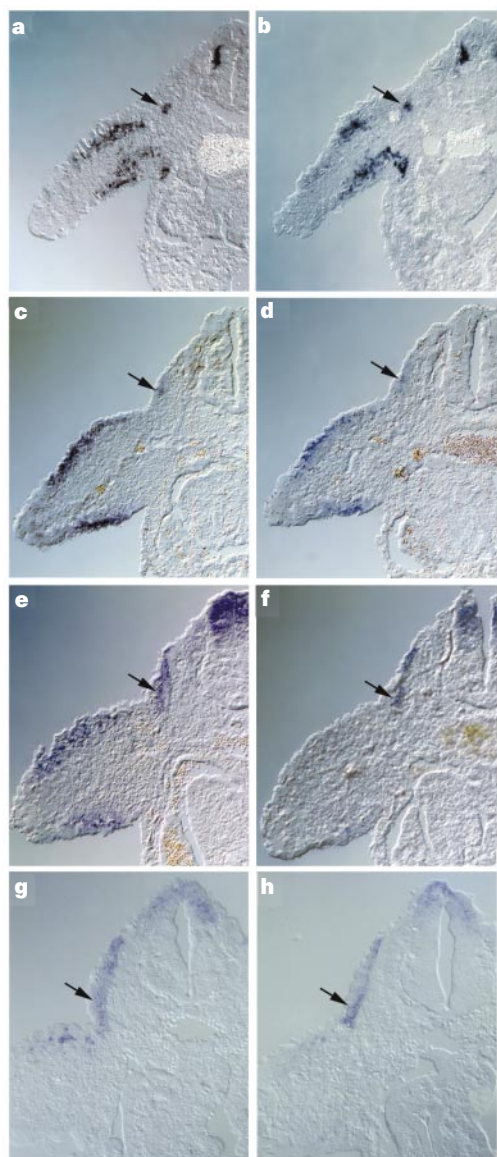


Figure 5 Gene expression in migrating limb myoblasts of *Mox2*^{-/-} embryos. Transverse sections of control (**a, c, e, g**) and mutant (**b, d, f, h**) embryos hybridized as whole-mount preparations with riboprobes specific for: *Lbx1* (**a, b**, *n* = 4). *c-Met* (**c, d**, *n* > 6) and *Pax3* (**e-h**, *n* > 6). **a-f**, E10.5 embryos; **g, h**, E10.0 embryos. Arrows indicate expression of each marker gene in the dermo-myotome. Note the specific downregulation of *Pax3* expression in limb buds of *Mox2*^{-/-} embryos (**f-h**).

MOX2-deficient embryos, expression of *MyoD* was normal in both the trunk and limb buds, but *Myf5*, although it was expressed normally in the trunk, was specifically downregulated in mutant limb buds (Fig. 6a, d). This molecular analysis did not reveal a defect in migration of limb myoblasts but showed that the *MOX2* homeoprotein is necessary for the maintenance of the normal gene expression profile of these cells. Furthermore, our data show that normal expression of the myogenic genes *MyoD* and *Myf5* in axial and appendicular muscles is differentially dependent on the function of the *MOX2* homeoprotein, which, although dispensable for the expression of both genes in trunk muscles, is necessary for normal expression of *Myf5* in limb muscles.

The association of *Pax3* and *Myf5* downregulation and the muscle defects observed in the limb buds of *Mox2*-deficient animals is consistent with genetic studies that have demonstrated that these genes are critical regulators of skeletal myogenesis; mouse embryos doubly homozygous for mutations in *Pax3* and *Myf5* fail to induce

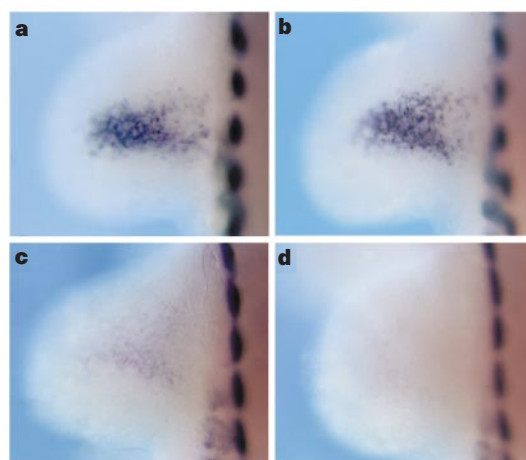


Figure 6 Differential expression of *MyoD* and *Myf5* genes in limb myoblasts of *Mox2*^{-/-} embryos. Dorsal aspects of whole-mount *in situ* hybridization analysis of E10.5 forelimb buds show that *MyoD* is expressed at similar levels in control (**a**, *n* = 4) and mutant (**b**, *n* = 4) embryos. In contrast, expression of *Myf5* is markedly downregulated in mutant forelimbs (**d**, *n* = 3) compared with controls (**c**, *n* = 3).

MyoD in the trunk and thus lack all axial skeletal muscles¹⁸. The normal *MyoD* expression detected in *Mox2*^{-/-} limb buds despite the downregulation of *Pax3* and *Myf5* indicates that the molecular mechanisms that activate *MyoD* expression are different in the limb and trunk muscle precursors. Furthermore, the severe defects of limb musculature in *Mox2*^{-/-} animals, notwithstanding high levels of *MyoD* expression, indicate that *MOX2* controls alternative *MyoD*-independent pathways that are necessary for the development of normal limb musculature. Thus, decreased expression of *c-Met* in *MOX2*-deficient limb myoblasts might contribute to the observed muscle defects. In support of this hypothesis, mice homozygous for a hypomorphic allele of the *c-Met* locus show defects in skeletal muscle differentiation¹⁹.

Two models could explain the role of *Mox2* in muscle development. According to the first model, *MOX2* functions within the myoblasts and its role could be mediated by *PAX3* and *MYF5*. This model is supported by: our expression studies on *Sp/Sp* mice which show that *Mox2* is expressed in the migrating myoblasts; the co-expression of *Mox2* with *Pax3* and genes for other myogenic factors (such as *MyoD* and presumably *Myf5*; see Supplementary information) in the myogenic lineage of limb buds; the dependence of the expression of the myogenic regulators *Pax3* and *Myf5* on normal *MOX2* function; and experiments on avian embryos which indicate that the early stages of limb muscle differentiation are cell autonomous and independent of the influence of lateral-plate-mesoderm-derived structures²⁰. According to the second model, the observed muscle defects may arise from abnormal development of connective tissue of the limbs in which *MOX2* function is required. This is consistent with expression of *Mox2* in limb-bud mesenchyme, which may give rise to connective tissue (Fig. 1) and experiments in avian embryos which have indicated that the characteristic muscle pattern elaborated during limb myogenesis can be controlled by connective tissues, such as tendons²⁰⁻²⁴. These models are not mutually exclusive and it is possible that different aspects of the phenotype of *MOX2*-deficient animals could be explained by either mechanism. In conclusion, our data show that the *MOX2* homeoprotein is a critical component of the regulatory framework that controls limb muscle development in vertebrate embryos. □

Methods

Targeted disruption of *Mox2*. Genomic clones were isolated from a 129/Sv library (Stratagene). We designed the targeting construct to replace a

2.6 kilobase (kb) *SacI*–*XhoI* fragment, containing the transcription and translation initiation sites, the first exon and part of the first intron, with the PGK–neo–polyA cassette. A PGK–thymidine kinase gene was also inserted to allow negative selection²⁵. A 6.5 kb *BamHI*–*SacI* fragment containing 5' genomic DNA was ligated to the *BamHI* site of the pPNT vector (a gift from V. Tybulewicz, NIMR), placing the 5' homology fragment downstream of the PGK–neo–polyA cassette (pKOMox2-BS6.5). The 3' homology (1.5 kb *XhoI*–*HindIII* fragment from intron 1) was inserted upstream of the PGK–neo–polyA cassette, between the *XhoI*–*NotI* sites of pKOMox2-BS6.5, to derive the final vector, pKOMox2. Electroporation of the linearized vector, selection and mouse blastocyst injections were done according to standard protocols. Ten (out of 150) clones contained the expected targeting event as determined by PCR and Southern blot analysis. Chimaeric males from four independent clones transmitted the mutation and generated heterozygous animals. Homozygous animals derived from these clones had identical phenotypes.

RNA analysis, histology, *in situ* hybridization and immunohistochemistry. Tail or yolk-sac DNA samples were genotyped with allele-specific PCR. We extracted RNA from whole embryos and various tissues using Trizol (Life Technologies) according to the manufacturer's recommendations. For northern blot analysis of total RNA (5 µg), we used radiolabelled antisense riboprobes derived from the 3' untranslated region of the *Mox2* gene. For histology, embryos or tissues were fixed in Bouin's fixative, dehydrated and embedded in paraffin. Serial sections (8 µm) were stained with haematoxylin and eosin.

Whole-mount *in situ* hybridization was done as described²⁶. For the double-labelling experiments, embryos were hybridized with *Pax3* antisense RNA, and subsequently incubated simultaneously with anti-DIG alkaline phosphatase conjugated antibody (Boehringer) and polyclonal rabbit anti-*Mox2* antibody. The *Mox2* signal was detected first by use of an anti-rabbit horseradish peroxidase-conjugated secondary antibody (DAKO), and peroxidase conversion of diaminobenzidine (Sigma) to an insoluble brown precipitate. The *Pax3* signal was then detected by incubation with NBT and BCIP (Boehringer) to generate a purple reaction product. For cryosectioning, embryos were postfixed in 4% paraformaldehyde, equilibrated in 30% sucrose and embedded in OCT. **Generation of myogenin–nlacZ mice.** We generated the myogenin–nlacZ (GNZ) transgene by fusing a 1.1 kb fragment from the *myogenin* gene promoter¹⁵ with a gene encoding nuclearly localized β-gal (a gift from E. Dzierzak, NIMR) in pBS KSII (Stratagene). The myogenin–nlacZ fragment was excised with *KpnI* and *XhoI*, and transgenic mice were generated by pronuclear injection. GNZ has the same expression pattern as previously described for GZ1092 (ref. 15). GNZ transgenic mice were crossed to *Mox2*^{+/-} animals and doubly heterozygous animals were intercrossed to derive *Mox2*^{-/-}; GNZ embryos. Mice carrying the GNZ transgene were identified by PCR amplification of genomic DNA. For β-gal histochemistry, embryos were fixed in 1X Mirsky's fixative (National Diagnostics) and processed as described¹⁵. Details of all methods are available upon request.

Received 4 January; accepted 22 April 1999.

- Christ, B., Jacob, H. J. & Jacob, M. Experimental analysis of the origin of the wing musculature in avian embryos. *Anat. Embryol. (Berl.)* **150**, 171–186 (1977).
- Buckingham, M. Making muscle in mammals. *Trends Genet.* **8**, 144–148 (1992).
- Cossu, G., Tajbakhsh, S. & Buckingham, M. How is myogenesis initiated in the embryo? *Trends Genet.* **12**, 218–223 (1996).
- Rudnicki, M. A. & Jaenisch, R. The MyoD family of transcription factors and skeletal myogenesis. *Bioessays* **17**, 203–209 (1995).
- Bober, E., Franz, T., Arnold, H. H., Gruss, P. & Tremblay, P. Pax-3 is required for the development of limb muscles: a possible role for the migration of dermomyotomal muscle progenitor cells. *Development* **120**, 603–612 (1994).
- Goulding, M., Lumsden, A. & Paquette, A. J. Regulation of Pax-3 expression in the dermomyotome and its role in muscle development. *Development* **120**, 957–971 (1994).
- Williams, B. A. & Ordahl, C. P. Pax-3 expression in segmental mesoderm marks early stages in myogenic cell specification. *Development* **120**, 785–796 (1994).
- Candia, A. F. *et al.* Mox-1 and Mox-2 define a novel homeobox gene subfamily and are differentially expressed during early mesodermal patterning in mouse embryos. *Development* **116**, 1123–1136 (1992).
- Candia, A. F. & Wright, C. V. Differential localization of Mox-1 and Mox-2 proteins indicates distinct roles during development. *Int. J. Dev. Biol.* **40**, 1179–1184 (1996).
- Daston, G., Lamar, E., Olivier, M. & Goulding, M. Pax-3 is necessary for migration but not differentiation of limb muscle precursors in the mouse. *Development* **122**, 1017–1027 (1996).
- Deconinck, A. E. *et al.* Utrophin-dystrophin-deficient mice as a model for Duchenne muscular dystrophy. *Cell* **90**, 717–727 (1997).
- Wachtler, E. & Jacob, M. Origin and development of the cranial skeletal muscles. *Bibl. Anat.* **29**, 24–46 (1986).
- Wright, W. E., Sassoon, D. A. & Lin, V. K. Myogenin, a factor regulating myogenesis, has a domain homologous to MyoD. *Cell* **56**, 607–617 (1989).
- Sassoon, D. *et al.* Expression of two myogenic regulatory factors myogenin and MyoD1 during mouse embryogenesis. *Nature* **341**, 303–307 (1989).

- Yee, S. P. & Rigby, P. W. The regulation of myogenin gene expression during the embryonic development of the mouse. *Genes Dev.* **7**, 1277–1289 (1993).
- Bladt, F., Riethmacher, D., Isenmann, S., Aguzzi, A. & Birchmeier, C. Essential role for the c-met receptor in the migration of myogenic precursor cells into the limb bud. *Nature* **376**, 768–771 (1995).
- Jagla, K. *et al.* Mouse Lbx1 and human LBX1 define a novel mammalian homeobox gene family related to the Drosophila ladybird genes. *Mech. Dev.* **53**, 345–356 (1995).
- Tajbakhsh, S., Rocancourt, D., Cossu, G. & Buckingham, M. Redefining the genetic hierarchies controlling skeletal myogenesis: Pax-3 and Myf-5 act upstream of MyoD. *Cell* **89**, 127–138 (1997).
- Maina, F. *et al.* Uncoupling of Grb2 from the Met receptor *in vivo* reveals complex roles in muscle development. *Cell* **87**, 531–542 (1996).
- Kardon, G. Muscle and tendon morphogenesis in the avian hind limb. *Development* **125**, 4019–4032 (1998).
- Chevallier, A., Kieny, M. & Mauger, A. Limb-somite relationship: origin of the limb musculature. *J. Embryol. Exp. Morphol.* **41**, 245–258 (1977).
- Jacob, H. & Christ, B. in *Teratology of the Limbs* (eds Merker, H. J., Nau, H. & Neubert, D.) 89–97 (Walter de Gruyter and Co., Berlin, 1980).
- Grim, M. & Wachtler, F. Muscle morphogenesis in the absence of myogenic cells. *Anat. Embryol. (Berl.)* **183**, 67–70 (1991).
- Yamamoto, M. *et al.* Coordinated expression of Hoxa-11 and Hoxa-13 during limb muscle patterning. *Development* **125**, 1325–1335 (1998).
- Mansour, S. L., Thomas, K. R. & Capocchi, M. R. Disruption of the proto-oncogene int-2 in mouse embryo-derived stem cells: a general strategy for targeting mutations to non-selectable genes. *Nature* **336**, 348–352 (1988).
- Henrique, D. *et al.* Expression of a Delta homologue in prospective neurons in the chick. *Nature* **375**, 787–790 (1995).

Supplementary information is available on Nature's World-Wide Web site (<http://www.nature.com>) or as paper copy from the London editorial office of Nature.

Acknowledgements. We thank S. Tajbakhsh for providing us with the *Splotch*^{2H} embryos; R. Balling, P. Gruss and K. Jagla for providing us with *Pax1*, *Pax3*, and *Lbx1* probes, respectively; T. Partridge for expert advice on histopathology; S. Hughes for useful discussions; H. Boyes for animal husbandry; and F. Costantini, in whose laboratory this work was initiated. This work was supported by the MRC.

Correspondence and request for materials should be addressed to V.P. (e-mail: v-pachni@nimr.mrc.ac.uk).

Plasmodium falciparum-infected erythrocytes modulate the maturation of dendritic cells

Britta C. Urban^{*}, David J. P. Ferguson[†], Arnab Pain^{*}, Nick Willcox^{*}, Magdalena Plebanski[‡], Jonathan M. Austyn[§] & David J. Roberts^{*¶}

^{*}Institute of Molecular Medicine, [†]Nuffield Department of Pathology, [‡]Molecular Immunology Group, [§]Nuffield Department of Surgery, [¶]National Blood Service—Oxford Centre, John Radcliffe Hospital, Headington, Oxford OX3 9DU, UK

The malaria parasite *Plasmodium falciparum* is one of the most successful human pathogens. Specific virulence factors remain poorly defined, although the adhesion of infected erythrocytes to the venular endothelium has been associated with some of the syndromes of severe disease¹. Immune responses cannot prevent the development of symptomatic infections throughout life, and clinical immunity to the disease develops only slowly during childhood. An understanding of the obstacles to the development of protective immunity is crucial for developing rational approaches to prevent the disease. Here we show that intact malaria-infected erythrocytes adhere to dendritic cells, inhibit the maturation of dendritic cells and subsequently reduce their capacity to stimulate T cells. These data demonstrate both a novel mechanism by which malaria parasites induce immune dysregulation and a functional role beyond endothelial adhesion for the adhesive phenotypes expressed at the surface of infected erythrocytes.

Specific immunity to malaria has been attributed either to the presence of cytotoxic lymphocytes that act against the parasite's liver stage of infection or to antibodies that react against blood-stage antigens^{2,3}. Antigenic diversity, clonal antigenic variation and T-cell antagonism may contribute to the parasite's evasion of the protective and parasitocidal host responses^{4–6}. Whatever the relative

importance of particular mechanisms and targets of acquired immunity, cytotoxic or helper T cells must be activated by antigen-presenting cells. Activated B cells, activated monocytes or dendritic cells can present antigen and stimulate activated T cells. However, dendritic cells are crucially important, as they can activate both memory and naive T cells, and hence initiate immune responses^{7,8}. We therefore investigated the outcome of the interaction between dendritic cells and malaria-infected erythrocytes.

Immature dendritic cells upregulate their surface expression of human leukocyte antigen (HLA) molecules, adhesion molecules and co-stimulatory molecules after they have been exposed to inflammatory stimuli such as lipopolysaccharide (LPS) or tumour necrosis factor- α (ref. 9). These mature dendritic cells are fully competent antigen-presenting and co-stimulatory cells. We found that when monocyte-derived dendritic cells were incubated with LPS for 48 h, the surface expression of HLA-DR, CD54, CD40, CD80, CD83 and CD86 antigens was readily increased (Fig. 1a). We also observed normal upregulation of the surface expression of HLA-DR, adhesion and co-stimulatory molecules when dendritic cells were incubated with LPS after exposure to intact uninfected erythrocytes, uninfected erythrocyte lysate, infected erythrocyte lysate, parasite-conditioned medium or a crude pigment preparation derived from infected erythrocytes (Fig. 1b, c, and data not shown). In marked contrast, dendritic cells exposed to intact infected erythrocytes did not show increased surface expression of these molecules in response to LPS (Fig. 1d). Inhibition of dendritic-cell maturation was induced after only 3 h of co-cultivation with infected erythrocytes at a ratio of infected erythrocytes to dendritic cells of 10:1 (data not shown).

The adhesion of infected erythrocytes to endothelial receptors is mediated, at least in part, by diverse parasite-derived proteins (PfEMP-1) that are expressed on the surface of infected erythrocytes and can undergo clonal antigenic variation¹⁰. Most parasite lines and clones adhere to CD36 and thrombospondin, and may adhere to one of a number of other receptors^{11–14}. The cloned parasite

line ITO/A4, used in our initial experiments, adheres to CD36, thrombospondin and CD54 (ICAM-1) (Fig. 2a). Both CD36 and CD54 are expressed on immature dendritic cells¹⁵. To test our hypothesis that the binding of infected erythrocytes to dendritic cells inhibits their maturation, we exposed dendritic cells to the parasite lines MC and ITO/C24, which adhere to CD36 and thrombospondin, and to the non-adherent parasite line T9/96, which does not express PfEMP-1 antigens at the red-cell surface (Fig. 2a, c, e, g)^{11,16}. Whereas the parasite lines MC and ITO/C24 inhibited the maturation of dendritic cells in a similar way to clone ITO/A4, the non-adherent line T9/96 did not prevent dendritic-cell maturation, even at a ratio of infected erythrocytes to dendritic cells of 100:1 (Fig. 2b, d, f, h).

We confirmed the adherence of ITO/A4-infected erythrocytes, but not T9/96-infected erythrocytes, to dendritic cells, using electron microscopy. Immature dendritic cells were incubated with ITO/A4- or T9/96-infected erythrocytes at a parasite-to-dendritic-cell ratio of 100:1 for 2 h. ITO/A4-infected erythrocytes were observed to be in intimate contact with immature dendritic cells, with cytoplasmic processes partially enclosing the parasites (Fig. 3a). The plasmalemma of the infected erythrocytes was in close apposition to the limiting membrane of the dendritic cell, particularly at the site of knobs (Fig. 3b). A similar apposition between parasitized erythrocytes and host cells is seen between infected red blood cells and endothelial cells¹⁷. In contrast, only a few infected erythrocytes of the T9/96 strain were associated with the dendritic cells (Fig. 3d). When quantified, we found 10 times more

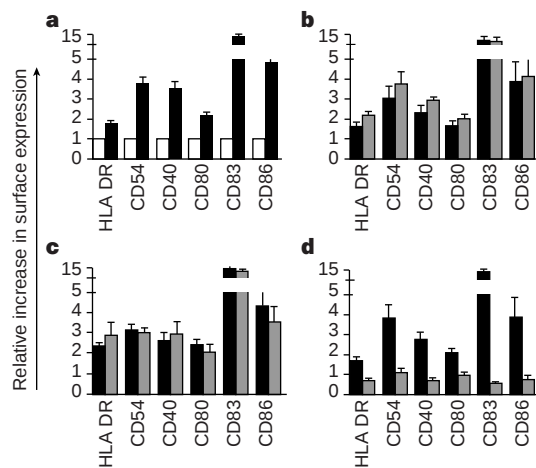


Figure 1 Expression of surface markers on dendritic cells in response to LPS. Increase in expression of surface molecules on: **a**, LPS-matured dendritic cells (black bars) compared with immature dendritic cells (white bars); **b**, dendritic cells matured with LPS with (hatched bars) and without (black bars) prior exposure to uninfected erythrocytes; **c**, dendritic cells matured with LPS with (hatched bars) and without (black bars) prior exposure to parasite lysate; **d**, dendritic cells matured with LPS with (hatched bars) and without (black bars) prior exposure to intact ITO/A4-infected erythrocytes. The differences in surface expression on dendritic cells exposed to intact infected erythrocytes compared with that on dendritic cells alone are statistically significant for all markers with $P < 0.01$ (Student's t -test). All experiments were repeated at least six times with dendritic cells from different donors. The relative increase in surface expression is expressed as the mean fluorescence intensity (MFI) of matured dendritic cells over the MFI of immature dendritic cells.

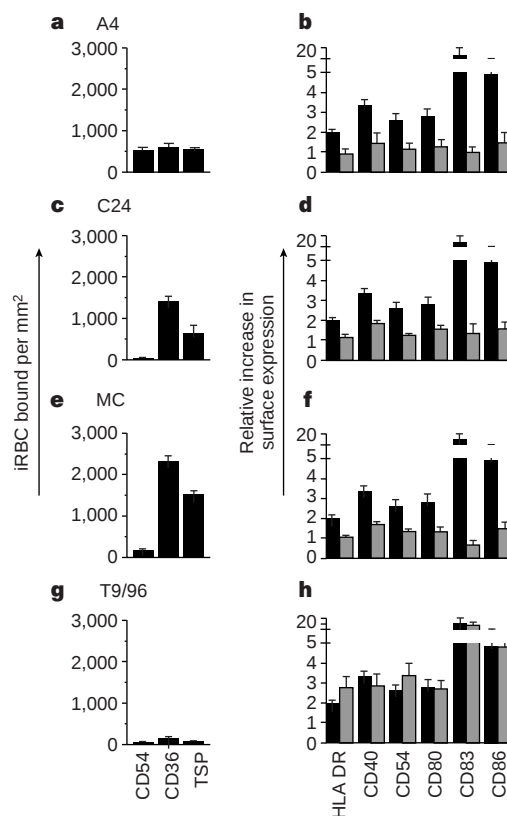


Figure 2 Only cytoadherent parasite lines inhibit dendritic-cell maturation. **a, c, e, g**, Absolute binding of infected erythrocytes (iRBC) of the respective parasite line to CD54, CD36 and thrombospondin (TSP); **b, d, f, h**, increase in surface expression of LPS-matured dendritic cells (black bars) compared with dendritic cells exposed to the parasite lines ITO/A4 (b), C24 (d), MC (f) and T9/96 (h) before maturation (hatched bars). The relative increase in surface expression on dendritic cells is expressed as the MFI of matured dendritic cells over the MFI of immature dendritic cells. All data represent the mean of three independent experiments.

ITO/A4-infected erythrocytes adhering to dendritic cells than T9/96-infected erythrocytes in thin sections of 100 dendritic cells. Furthermore, ingestion of intact ITO/A4-infected erythrocytes by dendritic cells was not observed during this time. Nevertheless, phagocytosis of parasite debris as revealed by the number of phagosomes containing pigment granules (Fig. 3c) was similar for dendritic cells incubated with erythrocytes infected with either the ITO/A4 or the T9/96 parasite strains.

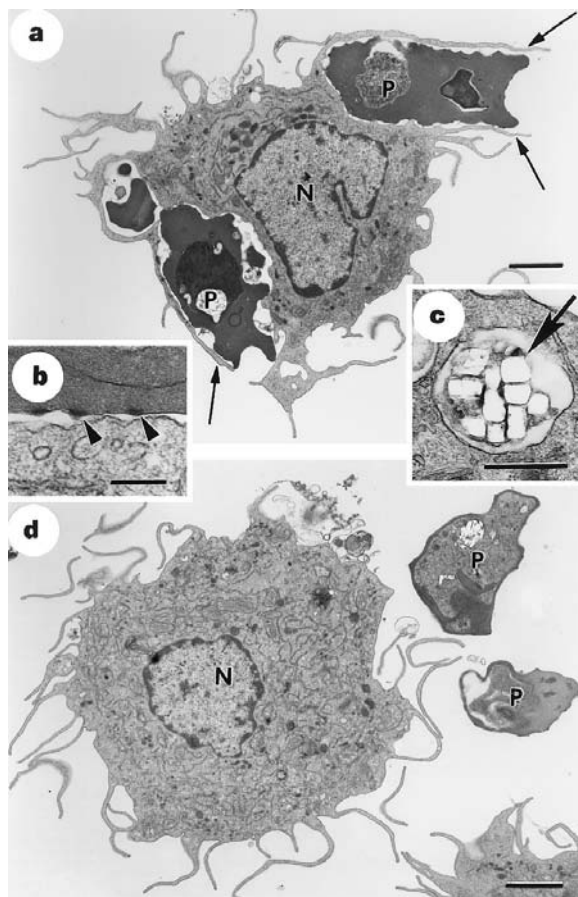


Figure 3 Transmission electron micrographs illustrating the interaction of dendritic cells with infected erythrocytes. Immature dendritic cells were mixed with adherent ITO/A4-infected erythrocytes (**a**) or with non-adherent T9/96-infected erythrocytes (**d**) for 2 h. Note the cell processes partially enclosing infected erythrocytes (arrows in **a**) and the close apposition of the limiting membranes of the infected erythrocytes and dendritic cells, particularly at the knobs (**b**, arrowhead). Within the dendritic-cell cytoplasm there are phagosomes containing characteristic pigment granules (**c**, arrows). N, dendritic-cell nucleus; P, infected erythrocyte. Scale bars: **a**, **d**, 2 µm; **b**, 200 nm; **c**, 500 nm.

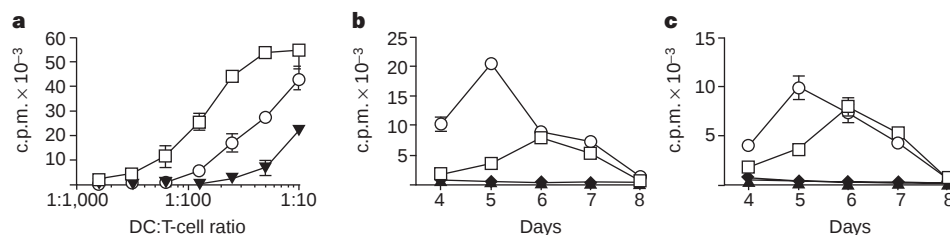


Figure 4 Dendritic cells (DC) exposed to intact infected erythrocytes are poor stimulators of primary T-cell responses. **a**, Allogeneic T-cell proliferation induced by immature dendritic cells (circles), LPS-matured dendritic cells (squares) and dendritic cells co-cultivated with intact ITO/A4-infected erythrocytes (inverted triangles) before maturation. **b**, **c**, Primary CD4⁺ T-cell responses to parasite lysate (**b**) and to KLH (**c**) induced by LPS-matured autologous dendritic cells (squares),

There was no significant difference in the proportion of cells undergoing early apoptosis (as assessed by the presence of annexin V on the cell surface but the absence of propidium iodide staining; annexinV+, PI-) or showing signs of necrosis (annexinV+, PI+) between control dendritic cells ($8.6 \pm 2\%$) or dendritic cells exposed to cytoadherent ($12 \pm 4\%$) or non-cytoadherent ($13 \pm 2.3\%$) parasite lines 72 h after exposure to parasites. These experiments demonstrate that dendritic-cell maturation is inhibited by the direct interaction with intact infected erythrocytes and is not due to a cytopathic effect of the parasite.

Mature dendritic cells are potent activators of T-cell proliferation^{7,8}. Dendritic cells that had matured after incubation with uninfected erythrocytes, a crude pigment preparation or a lysate of infected erythrocytes induced a similar degree of T-cell proliferation in a mixed leukocyte reaction (MLR) to that induced by control mature dendritic cells (data not shown). However, dendritic cells that were incubated with LPS after being exposed to intact infected erythrocytes from the ITO/A4 parasite line were strikingly less efficient in inducing T-cell proliferation than were mature dendritic cells or indeed immature dendritic cells (Fig. 4a). Similar results were obtained when primary responses of CD4-bearing (CD4⁺) T cells to parasite lysate or keyhole limpet haemocyanin (KLH) were analysed¹⁸. Dendritic cells pulsed with either antigen and then matured with LPS readily induced antigen-specific T-cell proliferation. Furthermore, mature dendritic cells alone induced a syngeneic MLR, but the kinetics of this response differed from the antigen-specific T-cell responses. However, exposure of dendritic cells to intact, cytoadherent infected erythrocytes before pulsing with KLH, parasite lysate or medium alone abolished not only the antigen-specific T-cell proliferation but also the syngeneic MLR (Fig. 4b, c), thereby confirming our observations in the allogeneic MLR.

Although secondary T-cell responses are less dependent on co-stimulation than are primary T-cell responses, we wondered whether dendritic cells exposed to cytoadherent infected erythrocytes also fail to stimulate antigen-specific T-cell lines and clones. We therefore generated CD4⁺ T-cell lines specific for KLH or for parasite lysate (Fig. 5a). Stimulation of these T-cell lines with dendritic cells that had been pulsed with KLH or parasite lysate, and then matured with LPS, induced T-cell proliferation. However, dendritic cells that had been exposed to intact infected erythrocytes before loading with either antigen and subsequent maturation with LPS failed to activate the respective T-cell lines (Fig. 5b). Finally, we examined whether parasite-exposed dendritic cells would support T-cell proliferation when loaded with an exogenous peptide, as these dendritic cells still express class II molecules of the major histocompatibility complex (MHC) and co-stimulatory molecules, albeit at a reduced level (see Fig. 1d). We therefore analysed a T-cell clone specific for the human acetylcholine receptor¹⁹. Dendritic cells pulsed with either the polypeptide or the restricting peptide induced proliferation of the T-cell clone. A similar degree of T-cell

and autologous dendritic cells co-cultivated with intact ITO/A4-infected erythrocytes (diamonds), pulsed with antigen (circles) or co-cultivated with intact ITO/A4-infected erythrocytes and then pulsed with antigen (triangles) before maturation. The data lines represented by diamonds and triangles overlaid one another. Data from one of three independent experiments are shown.

proliferation was observed when the dendritic cells were exposed to parasite lysate before the addition of the polypeptide or specific peptide, respectively. However, T-cell proliferation was abrogated when dendritic cells were exposed to intact infected erythrocytes before the polypeptide antigen was added (Fig. 5c). By contrast, dendritic cells exposed to intact infected erythrocytes, matured and loaded with the specific peptide, still induced a half-maximal response in the T-cell clone (Fig. 5d). Together, these results show that parasite-exposed dendritic cells fail to induce primary and secondary antigen-specific T-cell responses. Furthermore, erythrocytes infected with *P. falciparum* induce a significant defect in the processing of antigen by dendritic cells.

We conclude that the maturation of dendritic cells and their subsequent ability to activate T cells is profoundly inhibited by their interaction with intact infected erythrocytes. Non-adherent parasite lines, parasite debris and crude pigment do not modulate dendritic-cell function in this way. Our studies may provide one explanation for the clinical and experimental evidence of immune dysregulation seen during malaria infection, such as the impairment of the delayed-type hypersensitivity response to recall antigens and of the antibody response to vaccines^{20–22}.

It has been widely assumed that the adhesion of parasites has evolved solely to mediate sequestration of the parasites to endothelial cells in the peripheral tissues and so to reduce their destruction by splenic macrophages¹⁷. We now suggest that the adhesion of *P. falciparum*-infected erythrocytes to leukocytes is not a chance event, but confers a selective advantage upon the parasite: the acquisition of immune responses that kill parasites are delayed through modulation of the function of antigen-presenting cells. It is indeed intriguing that most *P. falciparum* wild isolates bind to CD36

and thrombospondin²³, although we cannot exclude the possibility that binding of parasites to other ligands on dendritic cells and leukocytes may have additional effects on the function of the immune system. Furthermore, in experimental models of animal malaria, strains that do not express variant antigens on the surface of infected erythrocytes, and thus do not cytoadhere, are avirulent and can induce immune responses that protect the animal against future heterologous challenge^{24–26}. These studies indeed proposed that the expression of variant antigens in the blood stages of malaria species may be involved in the suppression of immune responses²⁴. Finally, evidence is accumulating that all malaria species described express variant antigens, but not all cytoadhere to endothelium in their natural hosts^{27,28}. These observations and our data suggest that the expression of parasite-derived antigens at the surface of infected erythrocytes may have evolved to manipulate the immune system of their respective hosts. □

Methods

Cultivation of *P. falciparum*-infected red blood cells. Laboratory strains of *P. falciparum* were cultured in human red blood cells as previously described²⁹. The cytoadherent parasite lines ITO/A4 and ITO/C24 were clones isolated by micromanipulation from the ITO4 line, which is derived from a parasite isolate from Ituxi in Brazil. The cytoadherent parasite line Malayan Camp (MC) and the non-adherent parasite line T9/96 were both adapted to *in vitro* culture from parasites originally isolated in Thailand. All cultures were free from mycoplasma contamination. Infected erythrocytes were purified either by differential sedimentation in Plasmagel or through 65% Percoll, both of which gave a yield of more than 90% infected erythrocytes. Examination of a thin film revealed that more than 90% of infected erythrocytes were viable. Parasite lysate was obtained by three rounds of freezing and thawing of mature, infected red blood cells; parasite pigment was prepared as previously described³⁰. Parasite conditioned medium was the supernatant derived after 1×10^8 purified infected erythrocytes had been cultured in dendritic-cell medium for 24 h. All materials were from Sigma unless otherwise stated.

Binding of parasitized erythrocytes to purified proteins. This binding was measured as described previously⁵. Briefly, 2 μ l of a solution of thrombospondin (Gibco-BRL), purified CD36 or purified ICAM-Fc were adsorbed onto bacteriological, plastic plates. Mature, parasitized erythrocytes were suspended in binding medium and added to each dish. The erythrocytes were allowed to settle and then resuspended by gentle rotation every 10 min for 1 h. Non-adherent cells were removed, and the remaining cells were fixed and stained with Giemsa. Adherent parasitized cells were counted by light microscopy and the number of cells bound per mm^2 was corrected to binding at 2% haematocrit and 5% parasitaemia.

Generation of dendritic cells. Immature dendritic cells were derived from peripheral blood cells using standard procedures⁹. Briefly, monocytes were cultivated in RPMI 1640 supplemented with 2 mM glutamine, 50 $\mu\text{g ml}^{-1}$ kanamycin, 1% non-essential amino acids (Gibco-BRL), 10% human AB serum and 50 ng ml^{-1} each of interleukin 4 (specific activity $> 2 \times 10^6$ U per mg; PeproTech) and GM-CSF (specific activity $> 1 \times 10^7$ U per mg; Schering-Plough) for 6 days. Between days 6 and 9 of culture, non-adherent immature dendritic cells were collected and purified by depletion of contaminating lymphocytes using magnetic beads (Dyna) and anti-CD3 and anti-CD19 monoclonal antibodies (DAKO). For maturation assays, 1×10^6 purified dendritic cells were incubated in duplicate wells without antigen or with either 1×10^8 mature infected erythrocytes, 1×10^8 erythrocytes, parasite lysate corresponding to 1×10^8 infected erythrocytes or uninfected erythrocytes or a crude pigment preparation, for 20 h when not stated otherwise. Thereafter, dendritic cells were matured with 100 ng ml^{-1} LPS (from *Salmonella typhimurium*) for 48 h or left untreated as a control. Apoptosis of dendritic cells was evaluated with fluorescein isothiocyanate-annexin V/propidium iodide according to the manufacturers' recommendations (Boehringer-Mannheim).

Electron microscopy. One million purified immature dendritic cells were incubated for 2 h and for 12 h with 1×10^8 infected erythrocytes in 2 ml of dendritic-cell medium, collected and fixed with 2.5% glutaraldehyde/cacodylate buffer. Cells were post-fixed in osmium tetroxide, dehydrated and embedded in epoxy resin. Thin sections were stained with uranyl acetate and

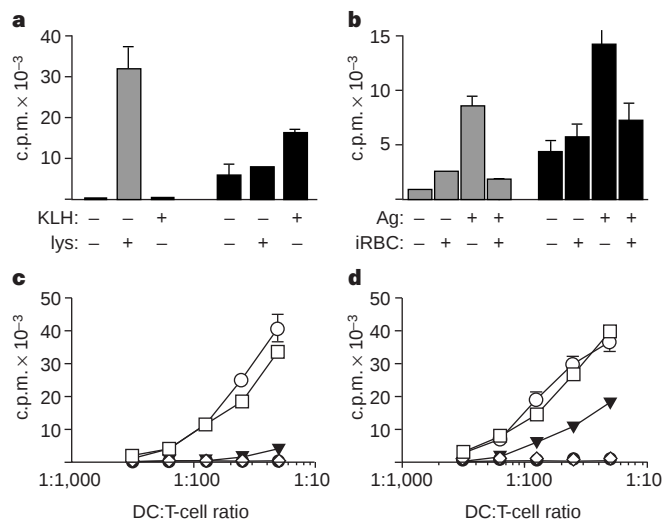


Figure 5 Dendritic cells exposed to intact infected erythrocytes are poor stimulators of secondary T-cell responses. **a**, Proliferation of T-cell lines specific for parasite lysate (hatched bars) or KLH (black bars) in response to medium, KLH or parasite lysate (lys) as indicated. **b**, Proliferation of T-cell lines specific for parasite lysate (hatched bars) or KLH (black bars) induced by autologous dendritic cells co-cultivated with intact infected erythrocytes (iRBC) and then pulsed with antigen (Ag) before maturation as indicated. **c**, **d**, Proliferation of the human acetylcholine-receptor (AChR)-specific T-cell clone TB-2 in response to 0.025 μM AChR α : amino acids 3–181 (**c**) or to 1 μM AChR α : amino acids 145–163 peptide (**d**). An increasing number of MHC class-II-matched dendritic cells alone (stars), pulsed with antigen (squares), co-cultivated with parasite lysate (black circles), co-cultivated with intact infected erythrocytes (open triangles), co-cultivated with parasite lysate and then pulsed with antigen (white circles), co-cultivated with intact infected erythrocytes and then pulsed with antigen (inverted filled triangles) before maturation with LPS were used as stimulator cells. The data lines represented by stars, filled circles and filled triangles overlap one another. Data from one of three independent experiments are shown.

lead citrate, then examined in a Jeol 1200EX electron microscope. The number of adherent and intracellular infected erythrocytes and the number of phagosomes containing pigment granules were counted in each sample in thin sections of 100 randomly selected dendritic cells.

Monoclonal antibodies and flow cytometry. The following monoclonal antibodies directed against the respective human surface markers were used: CD3, clone OKT3; HLA A,B,C, clone W6/32; CD14, clone Tük4; CD54, clone 6.5B5; CD19, clone HD37 (DAKO); CD36, clone 89; CD80, clone BB1; CD40, clone LOB7/6; CD86, clone BU63; HLA DR, clone BF-1 (Serotec); CD83, clone HB15a (a gift from T. F. Tedder)¹⁵. Dendritic-cell staining was performed as described¹⁵ and stained cells were analysed using a flow cytometer (Becton Dickinson).

T-cell proliferation assays. Total T cells (allogeneic MLR) or CD4⁺ T cells (primary T-cell responses) were purified using a Collect column (TCS). For the allogeneic MLR, dendritic cells were added in increasing numbers (156 to 10,000) to 1×10^5 T cells in triplicate and incubated for 5 days. T cells were pulsed with 0.5 μ Ci ³H-thymidine per well for the last 18 h of the culture. For antigen-specific T-cell responses, 1×10^6 dendritic cells were incubated with medium alone, with 1×10^8 infected erythrocytes or with parasite lysate corresponding to 1×10^8 infected erythrocytes for 18 h and then pulsed with antigen as specified (10 μ g ml⁻¹ parasite lysate, 30 μ g ml⁻¹ KLH, 0.025 μ M AChR- α : amino-acids 3–181, 1 μ M AChR- α : amino acids 145–163). Subsequently, dendritic cells were matured with LPS for a further 48 h and purified by sedimentation through Lymphoprep to remove parasite debris and dead cells. For primary T-cell responses, 1×10^5 dendritic cells were cultured with 1.5×10^6 autologous CD4⁺ T cells. From days 4 to 8 of culture, 50- μ l aliquots were taken in triplicate and pulsed with 0.5 μ Ci ³H-thymidine per well for 8 h¹⁸. For secondary T-cell responses of antigen-specific T-cell lines, 1×10^5 T cells were incubated in triplicate with 1×10^4 autologous dendritic cells for 4 days. For antigen-specific T-cell responses of clone TB-2, increasing numbers of MHC class-II-matched dendritic cells were incubated with 3×10^4 T cells for 72 h. Proliferation was measured in all assays by adding 0.5 μ Ci ³H-thymidine per well for the last 8 h of culture. Antigen-specific T-cell lines were generated in parallel with the primary T-cell responses from peripheral blood mononuclear cells according to standard procedures.

Received 26 February; accepted 22 April 1999.

- MacPherson, G. G., Warrell, M. J., White, N. J., Looareesuwan, S. & Warrell, D. A. Human cerebral malaria. A quantitative ultrastructural analysis of parasitized erythrocyte sequestration. *Am. J. Pathol.* **119**, 385–401 (1985).
- Hill, A. V. *et al.* Molecular analysis of the association of HLA-B53 and resistance to severe malaria. *Nature* **360**, 434–439 (1992).
- Marsh, K. *et al.* Antibodies to blood stage antigens of *Plasmodium falciparum* in rural Gambians and their relation to protection against infection. *Trans. R. Soc. Trop. Med. Hyg.* **83**, 293–303 (1989).
- Brown, K. N. Antigenic diversity, antigenic variation and merozoite surface protein 1. *Parasitologia* **35**, Suppl., 13–15 (1993).
- Roberts, D. J. *et al.* Rapid switching to multiple and adhesive phenotypes in malaria. *Nature* **357**, 689–692 (1992).
- Gilbert, S. C. *et al.* Association of malaria parasite population structure, HLA, and immunological antagonism. *Science* **279**, 1173–1177 (1998).
- Banchereau, J. & Steinman, R. M. Dendritic cells and the control of immunity. *Nature* **392**, 245–252 (1998).
- Austyn, J. M. Dendritic cells. *Curr. Opin. Hematol.* **5**, 3–15 (1998).
- Sallusto, S., Cella, M., Danieli, C. & Lanzavecchia, A. Dendritic cells use macropinocytosis and the mannose receptor to concentrate macromolecules in the major histocompatibility complex class II compartment: downregulation by cytokines and bacterial products. *J. Exp. Med.* **182**, 389–400 (1995).
- Smith, J. D. *et al.* Switches in expression of *Plasmodium falciparum* var genes correlate with changes in antigenic and cytoadherent phenotypes of infected erythrocytes. *Cell* **82**, 101–110 (1995).
- Roberts, D. D. *et al.* Thrombospondin binds falciparum malaria parasitized erythrocytes and may mediate cytoadherence. *Nature* **318**, 64–66 (1985).
- Barnwell, J. W. *et al.* A human 88-kD membrane glycoprotein (CD36) functions in vitro as a receptor for a cytoadherence ligand on *Plasmodium falciparum*-infected erythrocytes. *J. Clin. Invest.* **84**, 765–772 (1989).
- Berendt, A. R. *et al.* The binding site on ICAM-1 for *Plasmodium falciparum*-infected erythrocytes overlaps, but is distinct from, the LFA-1-binding site. *Cell* **68**, 71–81 (1992).
- Coppel, R. L., Cooke, B. M., Magowan, C. & Narla, M. Malaria and the erythrocyte. *Curr. Opin. Hematol.* **5**, 132–138 (1998).
- Zhou, L. J. & Tedder, T. F. Human blood dendritic cells selectively express CD83, a member of the immunoglobulin superfamily. *J. Immunol.* **154**, 3821–3835 (1995).
- Gardner, J. P., Pinches, R. A., Roberts, D. J. & Newbold, C. I. Variant antigens and endothelial receptor adhesion in *Plasmodium falciparum*. *Proc. Natl Acad. Sci. USA* **93**, 3503–3508 (1996).
- Udeinya, I. J., Schmidt, J. A., Aikawa, M., Miller, L. H. & Green, I. Falciparum malaria-infected erythrocytes specifically bind to cultured human endothelial cells. *Science* **213**, 555–557 (1981).
- Plebanski, M., Saunders, M., Burtles, S. S., Crowe, S. & Hooper, D. C. Primary and secondary human in vitro T cell responses to soluble antigens are mediated by subsets bearing different CD45 isoforms. *Immunology* **75**, 86–90 (1992).
- Nagykar, N. *et al.* A pathogenic role for the thymoma in myasthenia gravis. Auto-sensitization of IL-4-producing T-cell clones recognizing extracellular acetylcholine receptor epitopes presented by minority class II isotypes. *J. Clin. Invest.* **101**, 2268–2277 (1998).
- Williamson, B. A. & Greenwood, B. M. Impairment of the immune response to vaccination after acute malaria. *Lancet* **1**, 1328–1329 (1978).

- Greenwood, B. M., Bradley, A. K., Blakebrough, I. S. & Whittle, H. C. The immune response to a meningococcal polysaccharide vaccine in an African village. *Trans. R. Soc. Trop. Med. Hyg.* **74**, 340–346 (1980).
- Walsh, D. S., Looareesuwan, S., Vaninangonata, S., Virvan, C. & Webster, H. K. Cutaneous delayed-type hypersensitivity responsiveness in patients during and after *Plasmodium falciparum* and *Plasmodium vivax* infections. *Clin. Immunol. Immunopathol.* **77**, 89–94 (1995).
- Newbold, C. *et al.* Receptor-specific adhesion and clinical disease in *Plasmodium falciparum*. *Am. J. Trop. Med. Hyg.* **57**, 389–398 (1997).
- Howard, R. W. & Barnwell, J. W. Role of surface antigens on malaria-infected red blood cells in evasion of immunity. *Contemp. Top. Immunobiol.* **12**, 127–200 (1984).
- Langreth, G. E. & Peterson, E. Pathogenicity, stability, and immunogenicity of a knobless clone of *Plasmodium falciparum* in Colombian owl monkeys. *Infect. Immun.* **47**, 760–766 (1985).
- Gilks, C. E., Walliker, D. & Newbold, C. I. Relationships between sequestration, antigenic variation and chronic parasitism in *Plasmodium chabaudi chabaudi*—a rodent malaria model. *Parasite Immunol.* **12**, 45–64 (1990).
- Howard, R. J., Barnwell, J. W. & Kao, V. Antigenic variation of *Plasmodium knowlesi* malaria: identification of the variant antigen on infected erythrocytes. *Proc. Natl Acad. Sci. USA* **80**, 4129–4133 (1983).
- Udomsanpetch, R., Thanikkul, K., Pukrittayakamee, S. & White, N. J. Rosette formation by *Plasmodium vivax*. *Trans. R. Soc. Trop. Med. Hyg.* **89**, 635–637 (1995).
- Trager, W. & Jensen, J. B. Human malaria parasites in continuous culture. *Science* **193**, 673–675 (1976).
- Schwarzer, E., Turrini, F. & Arese, P. A luminescence method for the quantitative determination of phagocytosis of erythrocytes, of malaria-parasitized erythrocytes and of malarial pigment. *Br. J. Haematol.* **88**, 740–745 (1994).

Acknowledgements. We thank T. F. Tedder for the gift of anti-CD83 monoclonal antibody, D. Chao, P. Bahl, L. Corlett and K. Pulford for technical advice and C. Newbold, A. Vincent, B. C. Elford and S. Roberts for support and encouragement. D.J.P.F. was supported by the Wellcome Trust. D.J.R. is a Wellcome Trust Senior Fellow in Clinical Science and B.C.U. was supported by a Fellowship from the Deutsche Forschungsgemeinschaft.

Correspondence and requests for materials should be addressed to D.J.R. (e-mail: droberts@hammer.imm.ox.ac.uk).

Eph receptors and ephrins restrict cell intermingling and communication

Georg Mellitzer, Qiling Xu & David G. Wilkinson

Division of Developmental Neurobiology, National Institute for Medical Research, The Ridgeway, Mill Hill, London NW7 1AA, UK

Eph proteins are receptors with tyrosine-kinase activity which, with their ephrin ligands, mediate contact-dependent cell interactions¹ that are implicated in the repulsion mechanisms that guide migrating cells and neuronal growth cones to specific destinations^{2,3}. Ephrin-B proteins have conserved cytoplasmic tyrosine residues that are phosphorylated upon interaction with an EphB receptor^{4,5}, and may transduce signals that regulate a cellular response⁶. Because Eph receptors and ephrins have complementary expression in many tissues during embryogenesis⁷, bidirectional activation of Eph receptors and ephrin-B proteins could occur at interfaces of their expression domains, for example at segment boundaries in the vertebrate hindbrain. Previous work^{8,9} has implicated Eph receptors and ephrin-B proteins in the restriction of cell intermingling between hindbrain segments¹⁰. We therefore analysed whether complementary expression of Eph receptors and ephrins restricts cell intermingling, and whether this requires bidirectional or unidirectional signalling. Here we report that bidirectional but not unidirectional signalling restricts the intermingling of adjacent cell populations, whereas unidirectional activation is sufficient to restrict cell communication through gap junctions. These results reveal that Eph receptors and ephrins regulate two aspects of cell behaviour that can stabilize a distinct identity of adjacent cell populations.

Ephrins fall into two structural classes with different binding specificities: the ephrin-A ligands anchored with glycosyl phosphatidylinositol bind to the EphA class of receptors, whereas the transmembrane ephrin-B proteins bind to EphB receptors^{7,11}. An exception is EphA4, which binds ephrin-B2 as well as ephrin-A ligands⁷. Because truncated EphA4 or EphB receptors activate ephrin-B proteins as well as blocking Eph receptors, it is not possible to use these reagents to manipulate unidirectional and bidirectional signalling at interfaces of endogenous receptor and ephrin expression.

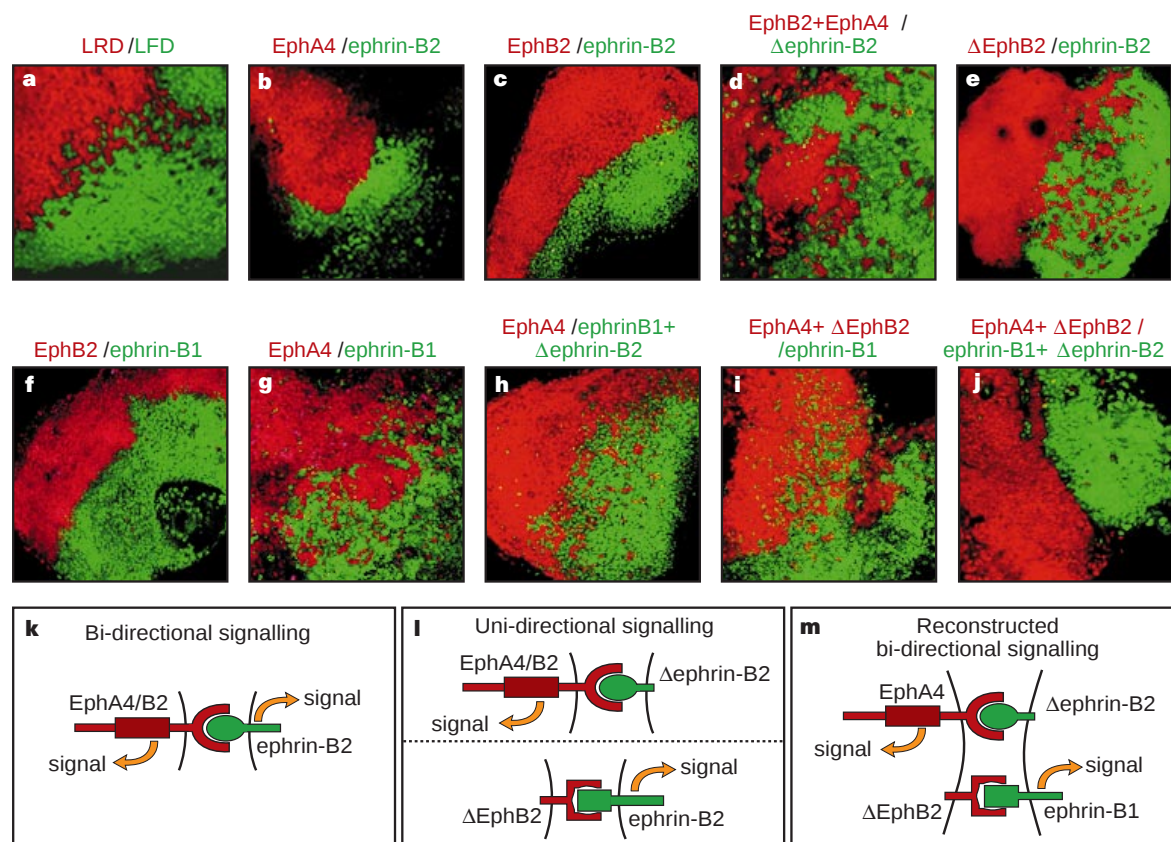


Figure 1 Bidirectional but not unidirectional signalling restricts cell intermingling. Animal caps that were labelled with rhodamine dextran (LRD) or fluorescein dextran (LFD) were juxtaposed and co-injected with RNA encoding Eph receptor or ephrin-B, respectively. After overnight culture, serial confocal sections of the fluorescent tracers were visualized. The following reagents were expressed in adjacent animal caps. **a**, Control assay: LRD/LFD only. **b**, **c**, Bidirectional signalling: **b**, EphA4/ephrin-B2; **c**, EphB2/ephrin-B2. **d**, **e**, Unidirectional signalling: **d**, EphA4 + EphB2/truncated ephrin-B2 (Δ ephrin-B2); **e**, truncated EphB2 (Δ EphB2)/ephrin-B2.

f–**j**, Controls and experiments to reconstruct bidirectional signalling: **f**, EphB2/ephrin-B1; **g**, EphA4/ephrin-B1; **h**, EphA4/ephrin-B1 + truncated ephrin-B2; **i**, EphA4 + truncated EphB2/ephrin-B1; **j**, reconstructed bidirectional signalling: EphA4 + truncated EphB2/ephrin-B1 + truncated ephrin-B2. **k**, Diagram depicting bidirectional activation of Eph receptor and ephrin-B protein. **l**, Diagram depicting unidirectional activation of Eph receptor by truncated ephrin-B2, or ephrin-B by truncated EphB2. **m**, Diagram depicting the combination of unidirectional activation used to reconstruct bidirectional signalling.

To circumvent this, we established an assay in which an Eph receptor and an ephrin are expressed in adjacent cell populations and the amount of cell intermingling determined. Zebrafish embryos at the one-cell stage are injected with fluorescent lineage tracer and then animal caps are dissected at the 1,000-cell stage. Upon juxtaposition of two animal caps, one labelled with rhodamine dextran and the other with fluorescein dextran, they rapidly adhere, and this aggregate is cultured overnight. Confocal microscopy revealed that intermingling occurs between uninjected control animal caps (Fig. 1a).

To test whether the interaction of Eph receptor and ephrin can restrict cell intermingling, we co-injected lineage tracer and RNA such that ephrin-B2 is expressed in one animal cap and EphB2 and/or EphA4 in the other. In all combinations, only a few cells crossed into the adjacent territory and a clear border was visible (Fig. 1b, c). In contrast, extensive cell intermingling occurred if the Eph receptor was omitted from one population, or ephrin from the other (data not shown). These results show that cell intermingling is restricted by the interaction of exogenous Eph receptor and ephrin-B2 at the interface of two cell populations.

To test whether the restriction of cell intermingling requires bi-directional activation (shown in Fig. 1k), we used truncated versions of ephrin-B2 and EphB2 lacking the intracellular domain, including tyrosine residues implicated in signal transduction, but which can act as ligands to activate phosphorylation of full-length Eph receptor or ephrin-B, respectively⁹. Juxtaposition of,

for example, cells expressing EphB2 with cells expressing truncated ephrin-B2 should therefore lead to unidirectional signalling into receptor-expressing cells (Fig. 1l). We found that after unidirectional signalling through Eph receptor (Fig. 1d) or ephrin-B (Fig. 1e), there is extensive cell intermingling between the two cell populations. To quantify the amount of cell intermingling, we counted the number of single labelled cells present in the adjacent territory in serial confocal sections (Fig. 2). Compared with uninjected controls, cell intermingling was significantly reduced by bidirectional signalling, but not by unidirectional signalling. These data indicate that unidirectional signalling in either direction is not sufficient, and bidirectional signalling is required to restrict cell intermingling.

A potential problem is raised by the discovery that Eph-receptor phosphorylation does not always correlate with a biological response, and that higher-order clustering of ephrin is required for functional activation of receptor¹². Because clustering could involve interaction of the intracellular domain of ephrins with cytoplasmic proteins¹³, it is possible that truncated ephrin does not fully activate Eph receptors. Therefore we tested whether truncated EphB2 and ephrin-B2 each elicit the relevant biological response. We devised a modified assay in which these reagents are used to reconstruct bidirectional signalling between cell populations from unidirectional signals. We took advantage of the finding that EphB2 binds ephrin-B1, whereas EphA4 binds ephrin-B2, but not ephrin-B1 (ref. 7). We found that restrictions to cell mixing

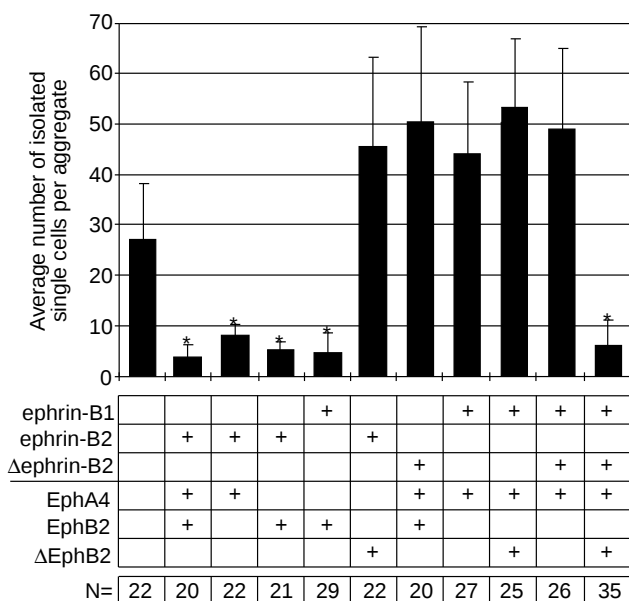


Figure 2 Quantification of cell intermingling. The bars indicate the average number of isolated single cells appearing in the adjacent territory for the indicated combinations of Eph receptors in one animal cap (below the line) and ephrin in the other (above the line). *N*, number of aggregates analysed. Intermingling is significantly reduced compared with uninjected controls ($P < 0.01$; indicated by asterisks) in combinations that give bidirectional activation of Eph receptor and ephrin-B. More intermingling compared with uninjected controls occurs ($P < 0.01$) when, for example, one cell population expresses full-length Eph receptor, and the other truncated ephrin. This could result from some autoactivation of full-length Eph receptor or ephrin-B that decreases cell-cell adhesion and thus increases the amount of cell movement.

correlate with these binding specificities, as EphB2 in combination with ephrin-B1 restricts cell intermingling, whereas EphA4 in combination with ephrin-B1 does not (Figs 1f, g, 2). Furthermore, intermingling occurs in the combination of (EphA4)/(ephrin-B1 + truncated ephrin-B2) (Figs 1h, 2), indicating that EphA4 does not activate ephrin-B1 indirectly through a heterodimer of ephrin-B1 and truncated ephrin-B2. Similarly, intermingling occurs in the combination of (ephrin-B1)/(EphA4 + truncated EphB2) (Figs 1i, 2), which argues against ephrin-B1 activating EphA4 through receptor heterodimerization. We could therefore reconstruct bidirectional signalling so that in one direction EphA4 is activated by truncated ephrin-B2, and in the other ephrin-B1 is activated by truncated EphB2 (Fig. 1m). In this situation, cell intermingling is restricted (Fig. 1j; quantification in Fig. 2). This result confirms that bidirectional signalling between two cell populations restricts their intermingling, but unidirectional signalling does not.

Correlations between sites of interaction between Eph receptors and ephrin in the hindbrain and somites^{2,3,14,15} and disruptions to cell communication through gap junctions^{16,17} led us to speculate that Eph receptors might regulate the formation of gap junctions. Gap junctions form by assembly of connexin proteins into intercellular channels that allow passage of molecules with a relative molecular mass below 1,200 ($M_r < 1.2K$)^{18,19}, and can be detected by the ability of Lucifer Yellow to diffuse through these channels. We juxtaposed one animal cap that was labelled with Lucifer Yellow (green in the confocal image), and another that was labelled with rhodamine dextran (red fluorescence). In the absence of co-injected reagents, Lucifer Yellow transfers into rhodamine dextran-labelled cells (the overlap leading to a yellow signal; Fig. 3a), indicating that gap junctions have formed between the cell populations. In contrast, when EphA4 or EphB2 were expressed in one animal cap and

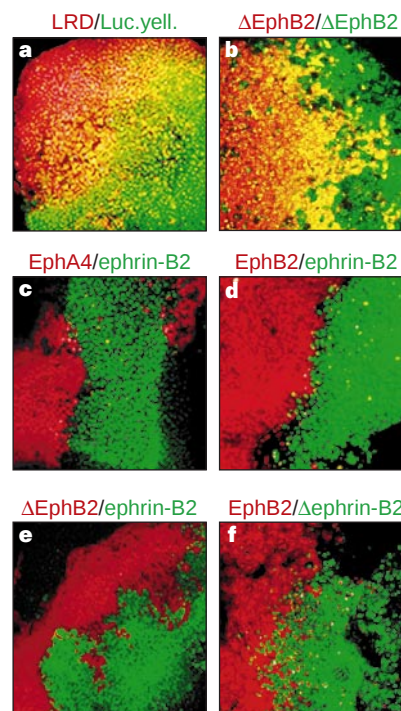


Figure 3 Activation of Eph receptor or ephrin blocks communication through gap junctions. Assays were carried out as described in Fig. 1, except that receptor-expressing embryos were labelled with LRD and ligand-expressing embryos were labelled with Lucifer Yellow (Luc.yell.), which gives a green signal on confocal microscopy. **a, b**, Control assays: **a**, LRD/Luc.yell.; **b**, truncated EphB2 (LRD)/truncated EphB2 (Luc.yell.). **c, d**, Bidirectional signalling: **c**, EphA4/ephrin-B2; **d**, EphB2/ephrin-B2. **e, f**, Unidirectional signalling: **e**, activation of ephrin-B2 by truncated EphB2; **f**, activation of EphB2 by truncated ephrin-B2.

ephrin-B2 in the other, Lucifer Yellow did not diffuse between the cell populations (Fig. 3c, d). This result indicates that bidirectional signalling blocks the formation of gap junctions, so we next tested the effect of unidirectional signalling. In control experiments, we found that expression of truncated EphB2 in adjacent animal caps did not restrict formation of gap junctions (Fig. 3b). In contrast, after unidirectional activation of ephrin-B2 by truncated EphB2 (Fig. 3e) or of EphB2 by truncated ephrin-B2 (Fig. 3f), Lucifer Yellow did not transfer into rhodamine-labelled cells, despite intermingling of the cell populations. Formation of gap junctions was also reduced after activation of ephrin-B2 with clustered soluble EphB1-Fc, but this reagent was less effective than membrane-bound truncated receptor (data not shown).

Based on these results, and evidence that Eph receptors and ephrins regulate repulsion or de-adhesion responses^{2,3}, we propose a model to account for the requirement for bidirectional signalling to restrict intermingling of adjacent cell populations, whereas unidirectional activation is sufficient to restrict communication by means of gap junctions (Fig. 4). At the interface of cells expressing Eph receptor and cells expressing ephrin-B, bidirectional activation leads to a mutual repulsion or de-adhesion that restricts the movement of each cell population into the other. In contrast, unidirectional signalling will repel one population, but the cells expressing truncated Eph receptor or ephrin are not themselves repelled and can invade adjacent territory, leading to intermingling. However, repulsion of only one of the two cell populations is sufficient to prevent stable cell-cell contacts required for assembly of gap junctions. After mosaic ectopic expression of truncated ephrin-B2 in zebrafish embryos, the expressing cells sort to the boundaries of rhombomeres r3/r5 (which express EphA4 and EphB receptors), indicating that unidirectional activation can restrict cells

to a specific region within segments⁹. This outcome, which is distinct from that observed with the animal-cap assay, could result from several differences between the systems. For example, in the hindbrain, stable boundaries are formed at which interactions between Eph receptors and ephrin occur, and r3/r5 cells expressing exogenous ephrin-B2 may sort because they have similar adhesive properties as boundary cells⁹.

There is much evidence that differential expression of cell adhesion molecules and the preferential association of cells with similar adhesive properties can establish and maintain organized cellular patterns during development^{20,21}. Although Eph receptor activation is required to prevent cell intermingling between hindbrain segments^{8,9}, this restriction may also require cell adhesion molecules²². Our findings reveal that complementary expression of Eph receptors and ephrins is sufficient to restrict cell intermingling, and that this can be accomplished without differential co-expression of exogenous adhesion molecules. However, Eph receptors and ephrins may regulate the function of cell adhesion molecules²³, leading to a de-adhesion of cells at boundaries. Alternatively, or in addition, activation may trigger cell repulsion responses involving localized collapse of the actin cytoskeleton²⁴. It is therefore likely that Eph receptors and ephrins have parallel or cooperative roles with cell adhesion systems in restricting cell intermingling during development.

Our results indicate that interfaces of endogenous Eph receptor and ephrin-B expression may also restrict formation of gap junctions. Communication by means of gap junctions has been implicated in

tissue patterning and the regulation of cell proliferation and differentiation^{18,19,25,26}, and it is believed to allow passage of regulatory molecules so that a coordinate response occurs in cells connected by gap junctions. In the hindbrain, rhombomere boundaries are barriers to the spread of signals that regulate regional identity, and this correlates with the absence of gap junctions²⁷. Similarly, formation of gap junctions is restricted at segment boundaries in insects²⁸. Several mechanisms could underlie such restrictions^{18,19}, including disassembly of connexins regulated by diffusible growth factors, and the interdependence of cell–cell adhesion and gap junction assembly. The regulation of gap junction formation by Eph receptors and ephrins enables a restriction of communication across boundaries, or even between intermingled cell populations. Thus, our findings indicate that interactions between Eph receptors and ephrin regulate two mechanisms—the restriction of cell intermingling and communication—that may stabilize a distinct identity or behaviour of adjacent cell populations. □

Methods

Animal cap assay for cell intermingling. Between 20 and 100 pg RNA encoding Eph receptor or ephrin (constructs described in ref. 9) was micro-injected into one-cell zebrafish embryos as described⁸, together with rhodamine dextran (LRD), fluorescein dextran (LFD) or Lucifer Yellow. At the 1,000-cell stage, embryos were dechorionated and animal caps were dissected. Upon juxtaposing animal caps, they adhere within several minutes to form an aggregate. Each aggregate was mounted under a coverslip and cultivated overnight in L15 medium containing 10% fetal calf serum. The aggregates were then fixed in 4% paraformaldehyde and equilibrated in 70% glycerol. Serial optical sections of the fluorescent tracers were visualized using a Leica confocal microscope. The images were displayed using NIH Image and processed with Adobe Photoshop software.

Received 1 February; accepted 5 May 1999.

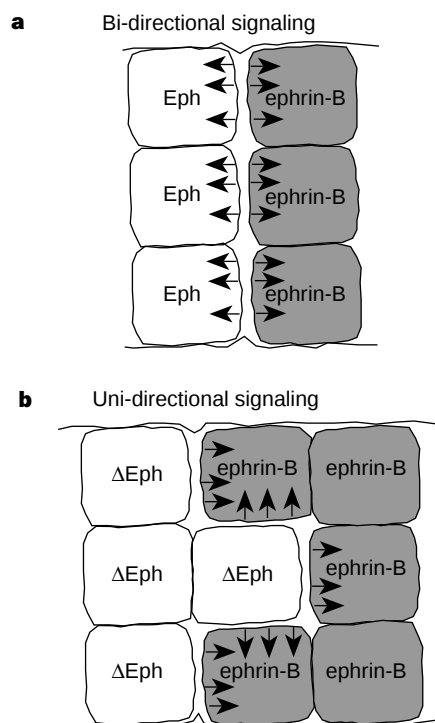


Figure 4 Model of restriction of cell intermingling and communication through gap junctions. **a**, Bidirectional activation leads to local de-adhesion and/or repulsion (indicated by arrows) in both the Eph-receptor- and ephrin-expressing cells at the interface of their expression domains. This prevents invasion of each cell population into the other, thus restricting cell intermingling. In addition, the absence of stable cell contacts (indicated by spaces between cells) disrupts the formation of gap junctions. **b**, Unidirectional activation leads to repulsion of cells expressing ephrin-B, but not of cells expressing truncated Eph receptor. This allows invasion of cells expressing truncated Eph receptor into ephrin-expressing territory, but repulsion of one cell population is sufficient to disrupt the formation of gap junctions. A similar situation occurs when cells expressing Eph-receptor and truncated ephrin-B are juxtaposed.

- Davis, S. et al. Ligands for EPH-related receptors that require membrane attachment or clustering for activity. *Science* **266**, 816–819 (1994).
- Flanagan, J. G. & Vanderhaeghe, P. The ephrins and Eph receptors in neural development. *Annu. Rev. Neurobiol.* **21**, 309–345 (1998).
- O'Leary, D. D. M. & Wilkinson, D. G. Eph receptors and ephrins in neural development. *Curr. Opin. Neurobiol.* **9**, 65–73 (1999).
- Holland, S. J. et al. Bidirectional signalling through the Eph-family receptor Nuk and its transmembrane ligands. *Nature* **383**, 722–725 (1996).
- Bruckner, K., Pasquale, E. B. & Klein, R. Tyrosine phosphorylation of transmembrane ligands for Eph receptors. *Science* **275**, 1640–1643 (1997).
- Henkemeyer, M. et al. Nuk controls pathfinding of commissural axons in the mammalian central nervous system. *Cell* **86**, 35–46 (1996).
- Gale, N. W. et al. Eph receptors and ligands comprise two major specificity subclasses, and are reciprocally compartmentalized during embryogenesis. *Neuron* **17**, 9–19 (1996).
- Xu, Q., Alldus, G., Holder, N. & Wilkinson, D. G. Expression of truncated Sek-1 receptor tyrosine kinase disrupts the segmental restriction of gene expression in the *Xenopus* and zebrafish hindbrain. *Development* **121**, 4005–4016 (1995).
- Xu, Q., Mellitzer, G., Robinson, V. & Wilkinson, D. G. *In vivo* cell sorting in complementary segmental domains mediated by Eph receptors and ephrins. *Nature* **399**, 267–271 (1999).
- Fraser, S., Keynes, R. & Lumsden, A. Segmentation in the chick embryo hindbrain is defined by cell lineage restrictions. *Nature* **344**, 431–435 (1990).
- Eph Nomenclature Committee. Unified nomenclature for Eph family receptors and their ligands, the ephrins. *Cell* **90**, 403–404 (1997).
- Stein, E. et al. Eph receptors discriminate specific ligand oligomers to determine alternative signaling complexes, attachment, and assembly responses. *Genes Dev.* **12**, 667–678 (1998).
- Torres, R. et al. PDZ domains bind, cluster, and synaptically colocalize with Eph receptors and their ephrin ligands. *Neuron* **21**, 1453–1463 (1998).
- Wang, H. U. & Anderson, D. J. Eph family transmembrane ligands can mediate repulsive guidance of trunk neural crest migration and motor axon outgrowth. *Neuron* **18**, 383–396 (1997).
- Krull, C. E. et al. Interactions of Eph-related receptors and ligands confer rostrocaudal pattern to trunk neural crest migration. *Curr. Biol.* **7**, 571–580 (1997).
- Martinez, S., Geijjo, E., Sanchez-Vives, M. V., Puelles, L. & Gallego, R. Reduced junctional permeability at interrhombomeric boundaries. *Development* **116**, 1069–1076 (1992).
- Bagnall, K. M., Sanders, E. J. & Berdan, R. C. Communication compartments in the axial mesoderm of the chick embryo. *Anat. Embryol.* **186**, 195–204 (1992).
- Kumar, N. H. & Gitula, N. B. The gap junction communication channel. *Cell* **84**, 381–388 (1996).
- Bruzzone, R., White, T. W. & Paul, D. L. Connections with connexins: the molecular basis of direct intercellular signaling. *Eur. J. Biochem.* **238**, 1–27 (1996).
- Takeichi, M. Cadherin cell adhesion receptors as a morphogenetic regulator. *Science* **251**, 1451–1455 (1991).
- Gumbiner, B. M. Cell adhesion: the molecular basis of tissue architecture and morphogenesis. *Cell* **84**, 345–357 (1996).
- Wizenmann, A. & Lumsden, A. Segregation of rhombomeres by differential chemoaffinity. *Mol. Cell. Neurosci.* **9**, 448–459 (1997).
- Zisch, A. H. et al. Tyrosine phosphorylation of L1 family adhesion molecules: implication of the Eph kinase Csk5. *J. Neurosci. Res.* **47**, 655–665 (1997).
- Bruckner, K. & Klein, R. Signaling by Eph receptors and their ephrin ligands. *Curr. Opin. Neurobiol.* **8**, 375–382 (1998).

25. Lo, C. W. The role of gap junction membrane channels in development. *J. Bioenerg. Biomemb.* **28**, 379–385 (1996).
26. Simon, A. M. & Goodenough, D. A. Diverse functions of vertebrate gap junctions. *Trends Cell Biol.* **8**, 477–483 (1998).
27. Martinez, S., Marin, F., Nieto, M. A. & Puellas, L. Induction of ectopic engrailed expression and fate change in avian rhombomeres: intersegmental boundaries as barriers. *Mech. Dev.* **51**, 289–303 (1995).
28. Warner, A. E. & Lawrence, P. A. Permeability of gap junctions at the segmental border in insect epidermis. *Cell* **28**, 243–252 (1982).

Acknowledgements. We thank M. Henkemeyer, N. Gale and G. Yancopoulos for reagents, and J.-P. Vincent and R. Krumlauf for helpful discussions and comments on the manuscript. This work was supported by an EMBO Postdoctoral Fellowship to G.M., the MRC and a European Commission Biotechnology grant.

Correspondence and requests for materials should be addressed to D.G.W. (e-mail: d-wilkin@nimr.mrc.ac.uk).

DNA-dependent protein kinase is not required for the p53-dependent response to DNA damage

Gretchen S. Jimenez^{*†}, Fredrik Bryntesson^{†‡}, Maria I. Torres-Arzuayus[§], Anne Priestley^{||}, Michelle Beeche^{*}, Shin'ichi Saito[¶], Kazuyasu Sakaguchi[¶], Ettore Appella[¶], Penny A. Jeggo^{||}, Guillermo E. Taccioli[§], Geoffrey M. Wahl^{*†} & Michael Hubank[†]

^{*} Gene Expression Laboratory, The Salk Institute, La Jolla, California 92037, USA

[†] Trafford Centre, University of Sussex, Brighton, Sussex BN1 9RY, UK

[§] Department of Microbiology, Boston University School of Medicine, Boston, Massachusetts 02118, USA

^{||} MRC Cell Mutation Unit, University of Sussex, Brighton, Sussex BN1 9RR, UK

[¶] Laboratory of Cell Biology, National Cancer Institute, National Institute of Health, Bethesda, Maryland 20892, USA

[†] These authors contributed equally to this work

Damage to DNA in the cell activates the tumour-suppressor protein p53 (ref. 1), and failure of this activation leads to genetic instability and a predisposition to cancer. It is therefore crucial to understand the signal transduction mechanisms that connect DNA damage with p53 activation. The enzyme known as DNA-dependent protein kinase (DNA-PK) has been proposed to be an essential activator of p53 (refs 2, 3), but the evidence for its involvement in this pathway is controversial^{3,4}. We now show that the p53 response is fully functional in primary mouse embryonic fibroblasts lacking DNA-PK: irradiation-induced DNA damage in these defective fibroblasts induces a normal response of p53 accumulation, phosphorylation of a p53 serine residue at position 15, nuclear localization and binding to DNA of p53. The upregulation of p53-target genes and cell-cycle arrest also occur normally. The DNA-PK-deficient cell line SCGR11 contains a homozygous mutation in the DNA-binding domain of p53, which may explain the defective response by p53 reported in this line³. Our results indicate that DNA-PK activity is not required for cells to mount a p53-dependent response to DNA damage.

DNA-PK^{-/-} primary mouse embryonic fibroblasts (MEFs) were derived from a mouse harbouring a targeted disruption of the catalytic subunit of DNA-PK which causes loss of the kinase activity⁵. The phenotype of these MEFs is radiosensitive, as are the phenotypes of SCID (severe combined immunodeficient) mice and of high-passage-number transformed lines derived from these mice, such as SCGR11 (Fig. 1a). Although it is unclear whether there is a residual DNA-PK function in SCID-derived cells, DNA-PK^{-/-} MEFs have no DNA-PK activity or any detectable DNA-PK protein by western blotting⁵. If DNA-PK were required for p53-mediated cell-cycle arrest in response to DNA damage, then DNA damage should not induce a cell-cycle arrest in DNA-PK^{-/-} cells. However, like wild-type cells, irradiated DNA-PK^{-/-} MEFs accumulate in G1

phase as a result of cell-cycle arrest (Fig. 1b); we also found that irradiated DNA-PK^{-/-} MEFs have a higher G1/S ratio than DNA-PK^{+/+} or +/- cells irradiated with the same dose, probably because of the persistence of unrepaired double-stranded DNA breaks. DNA-PK^{-/-} MEFs also undergo arrest of the cell cycle in response to the antimetabolite PALA (*n*-(phosphonacetyl)-L-aspartate) or the microtubule-depolymerizing agent colcemid (data not shown). In contrast, p53^{-/-} MEFs and SCGR11 cells continue to cycle after DNA damage (Fig. 1b).

We next investigated the stabilization of p53 and its mobilization to the nucleus after different types of DNA damage. Using indirect immunofluorescence to track the subcellular localization of p53 in wild-type and DNA-PK^{-/-} MEFs, we found similar rates of nuclear accumulation and a comparable duration of nuclear localization after γ -irradiation or treatment with actinomycin D in both cell types (Fig. 2A, and data not shown). By contrast, SCGR11 cells contained large amounts of nuclear p53 even in untreated cells, which is indicative of the presence of mutant, stable p53 (Fig. 2A).

p53 is phosphorylated at highly conserved amino-terminal serine residues in response to DNA damage^{2,6}. Phosphorylation of p53 at serine 15 reportedly blocks binding of the p53 inhibitor MDM2, preventing MDM2-mediated degradation of p53. This stabilized p53 protein can then transactivate its target genes^{2,6}. The structurally related kinases ATM, ATR and DNA-PK all phosphorylate p53 at Ser 15 *in vitro*^{7,8}. However, phosphorylation of Ser 15 and accumulation of p53, monitored by immunoprecipitation after either ultraviolet or γ -irradiation, occurred to the same extent in DNA-PK^{-/-} and wild-type MEFs (Fig. 2B, C, and data not shown). We conclude that DNA-PK is therefore dispensable for the nuclear accumulation of p53, phosphorylation of Ser 15, and stabilization of p53 that occur in response to DNA damage.

Once in the nucleus, p53 binds to specific DNA sequences in promoters and enhancers of genes that are involved in cell-cycle arrest and apoptosis¹. It has been proposed that p53 must be

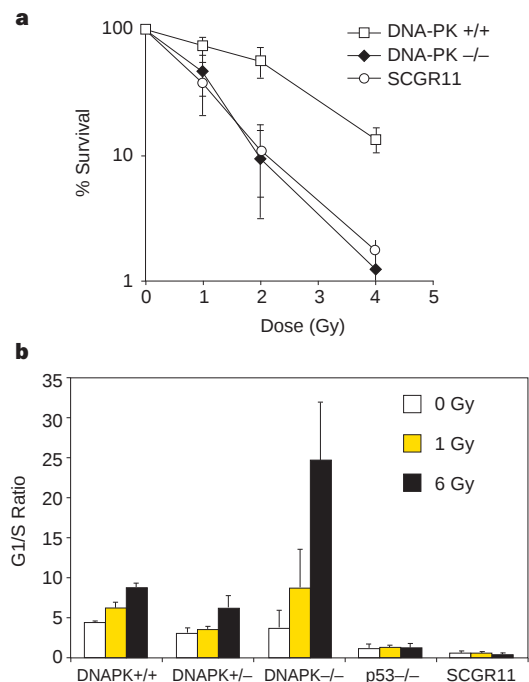


Figure 1 Cell-cycle arrest response of DNA-PK^{-/-} MEFs following irradiation. **a**, DNA-PK^{-/-} cells show a level of radiosensitivity that is identical to that of SCGR11 cells. The results represent the mean $\pm$ s.d. of 3 independent experiments. **b**, Using dual-parameter microfluorimetry, G1, S or G2/M phase was assessed on >10,000 collected cells per sample. Fluorescence-activated cell sorting (FACS) analysis was performed at least twice for each treatment and the results represent the mean of the G1/S phase ratio $\pm$ s.d.

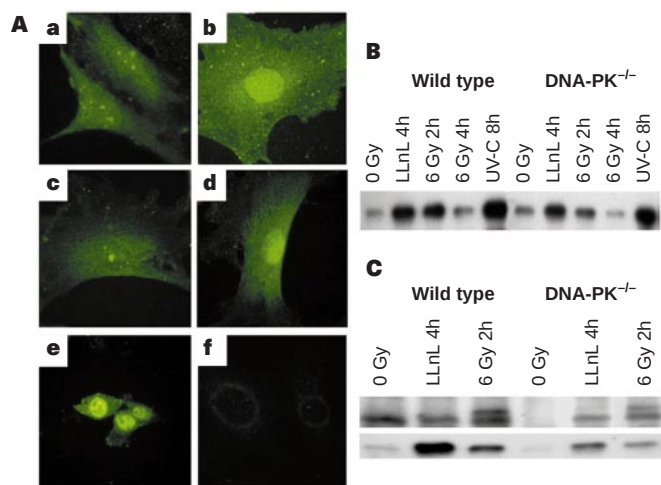


Figure 2 Accumulation and localization of p53 protein. **A**, p53 localization in untreated DNA-PK^{-/-} MEFs (**a**), 4 h after 6 Gy γ -irradiation (**b**), untreated wild-type MEFs (**c**), 4 h after 6 Gy γ -irradiation (**d**), untreated SCGR11 cells (**e**), and p53^{-/-} MEFs treated with actinomycin D for 4 h (**f**). **B**, Immunoprecipitation of p53 protein from wild-type and DNA-PK^{-/-} MEFs. **C**, Immunoprecipitation followed by western blotting with anti-phosphoserine-15 antibody (top, upper band; 52K), and control blotted with Ab421 (bottom). The anti-phosphoserine-15 blot contains a nonspecific band at 50K.

modified by DNA-PK to bind efficiently to DNA³, but we find that endogenous p53 binds to a specific DNA-target oligomer equally well in extracts prepared from irradiated DNA-PK^{-/-}, +/- or +/+ cells (Fig. 3a). In contrast, p53 in nuclear extracts from irradiated SCGR11 cells does not bind to this p53 consensus site. We next tested whether DNA-PK activity affected the transactivation of p53 target genes. Analysis using the polymerase chain reaction with reverse transcription (RT-PCR) revealed that the p53 target genes *p21*, *GADD45* and *MDM2* were all upregulated in irradiated DNA-PK^{-/-} primary MEFs to the same extent as in wild-type MEFs (Fig. 3b). We conclude that the entire p53 response is intact in DNA-PK^{-/-} MEFs.

In contrast to DNA-PK^{-/-} MEFs, the mutant line SCGR11 showed no detectable p53 response in any of our experiments; also, there is no p53-mediated upregulation of p21 in response to irradiation³. We examined the p53 transactivational response in a range of transformed fibroblast lines carrying mutations in DNA-PK or in its Ku80 subunit^{4,9}. p21 and, to a lesser extent, *MDM2*, were induced by γ -irradiation in all cell lines, regardless of their DNA-PK phenotype, with the exception of the SCGR11 cell line (Fig. 3b). Sequencing of p53 in SCGR11 identified a homozygous T-to-G transversion at nucleotide 572 in the DNA-binding region of p53, at which an arginine is substituted for a leucine at codon 191. This corresponds to a mutation in human p53 (at codon 194) that has been identified in leukaemias, lymphomas and cancers of the lung, breast and oesophagus¹⁰. The p53 sequence in DNA-PK^{-/-} MEFs was wild-type.

The mutant p53 found in SCGR11 cells is stable, nuclear, and transcriptionally inactive (Figs 2, 3); these cells can support transcriptional activation when transfected with wild-type p53 (D. Lane, personal communication). These results cast doubt upon the idea that DNA-PK is necessary for p53 activation in these cells³, although Woo *et al.* detect specific binding of p53 from both wild-type and SCGR11 cells to DNA, apparently in a DNA-PK dependent manner³. This discrepancy with our findings may be explained if the pAb421 antibody needed for binding in these experiments can stabilize a non-physiological interaction between an irradiation-modified form of wild-type or mutant p53 and its DNA target sequence. In the case of SCGR11, this modification must induce a conformational change in its non-binding mutant p53 to a form

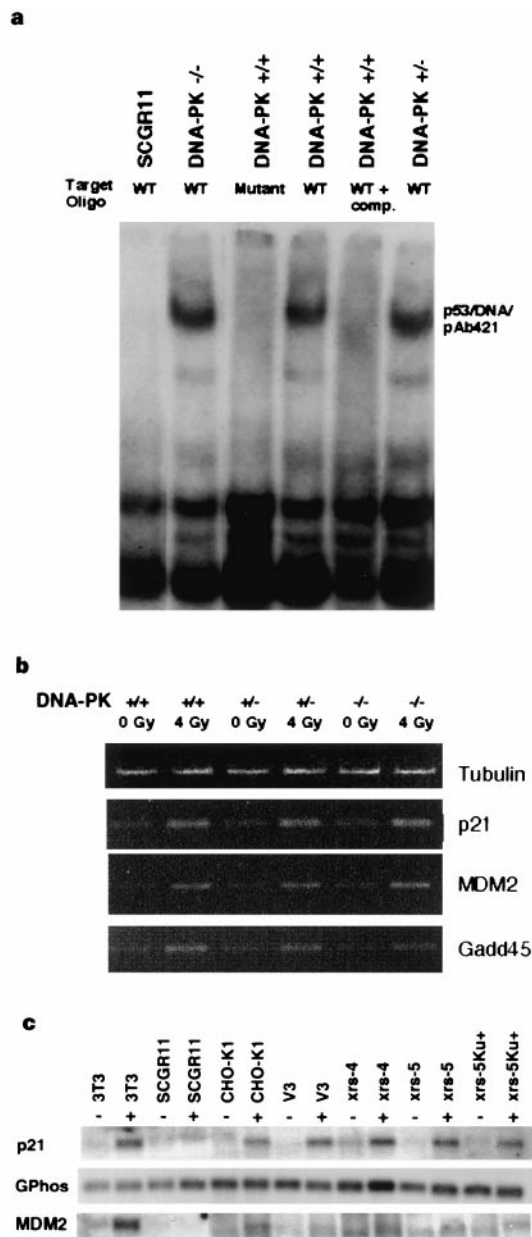


Figure 3 p53 binding and transactivation in DNA-PK^{-/-} MEFs. **a**, Gel-shift analysis of p53 binding to a p53-specific oligonucleotide template. All reactions were incubated with pAb421. **b**, Activation of p53-dependent target genes by γ -irradiation. Semi-quantitative RT-PCR was performed using tubulin as a control. *p21*, *MDM2* and *GADD45* were upregulated in all groups in response to irradiation. **c**, Northern blot analysis of irradiated transformed lines.

capable of binding DNA. The positioning of the mutated amino acid away from the DNA-binding surface of the molecule would allow this. However, although this raises interesting questions about the way in which p53 phosphorylation influences binding activity, it is clear that even if such DNA-PK-dependent modification occurs *in vivo*, it is not required for p53 activation.

There is other evidence against a role for DNA-PK in the activation of p53. If p53 is downstream of DNA-PK in the signalling pathway, there should be no requirement for selection of a p53 mutation in SCGR11. Neither DNA-PK- nor Ku-deficient mice have a p53^{-/-} phenotype, and there does not appear to be a failure of DNA-damage-induced cell-cycle arrest in these mice^{5,11-13}. Furthermore, cells deficient in ATM kinase, which still contain functional DNA-PK, do not arrest effectively after DNA damage, arguing against a redundant role for DNA-PK in the p53 response¹⁶. It

is more likely that the ATM and ATR kinases are better candidates for activating p53, although DNA-PK may operate differently in humans. We conclude that DNA-PK does not act upstream of p53 in the DNA-damage-response pathway. □

Methods

MEFs, cell lines and reagents. DNA-PK^{-/-}, +/- and +/+ primary MEFs were genotyped and cultured for a maximum of seven passages in RPMI medium with 10% FCS, 1 mM 2-mercaptoethanol and antibiotics, or in DMEM medium with 10% FCS and antibiotics. DNA-PK-defective fibroblast lines, xrs4, xrs5, xrs6 (hamster Ku80 mutants), V3 (hamster DNA-PK mutant) and SCRG11 (mouse DNA-PK mutant), together with controls xrs5+Ku (xrs5 complemented with wild-type Ku80), wild-type cells CHO-K1 (hamster), p53^{-/-} MEFs and 3T3 (mouse) cells, were grown in DMEM with 10% FCS, L-glutamine, NEAA and antibiotics. Cells were γ-irradiated at room temperature at a dose rate of 2 Gy min⁻¹ with a ⁶⁰Co γ-irradiator UV-C (25 J m⁻²) radiation was administered at a dose rate of 4.4 J m⁻² s⁻¹. Calpain inhibitor I (LLnL; sigma) was used at 20 μM for 4 h. Actinomycin D (Sigma) was used at 300 ng ml⁻¹.

Survival. MEFs that had been γ-irradiated were plated on a feeder layer of lethally irradiated cells and the surviving colonies were counted ten days later.

Cell cycle analysis. Cells from at least two independently derived embryos were used for DNA-PK^{+/+}, +/- and -/- MEFs. Log-phase cells were irradiated with 1 Gy or 6 Gy γ-radiation and incubated for 24 h. Cells were fixed in 70% ethanol and double stained with BrDU-PI as described¹⁷.

Immunofluorescence. Cells were irradiated on their coverslips, fixed in 1:1 acetone:methanol and stained with anti-p53 antibody Ab421 (CalBiochem Ab-1), followed by FITC-conjugated goat anti-mouse secondary (Boehringer Mannheim) and Hoescht 33258 (Sigma). Images were captured using a 60× objective and the Deltavision system.

Immunoprecipitation. Cell extracts were prepared and western blotted as described¹⁸, except that mouse p53 was collected by centrifugation with pAb421-agarose beads. Affinity-purified rabbit polyclonal antibodies against a human peptide containing phosphorylated serine 15 were prepared as described¹⁸.

Gel-shift assay. Nuclear extracts were prepared 4 h after 6 Gy γ-irradiation¹⁹. The final DNA-binding reactions contained 100 mM KCl, 20 mM HEPES, pH 7.6, 1 mM DTT, 10% glycerol, 100 ng salmon sperm DNA, and 0.1 ng γ-³²P end-labelled p53 consensus or mutant double-stranded oligonucleotide (Santa Cruz Biotechnology). Samples were incubated for 30 min at room temperature and then for a further 15 min after addition of 0.5 μg Ab421 to supershift p53-DNA complexes. 2 ng (20-fold excess) of unlabelled consensus oligonucleotide was added as a competitor where indicated. Samples were run on a 5% Tris-borate-EDTA polyacrylamide gel.

RT-PCR. DNA-PK -/-, +/- and +/+ MEFs were irradiated with 4 Gy and incubated at 37°C for 4 h. Total cellular RNA was prepared using the TRIZOL (Gibco) method. Residual genomic DNA was removed with DNase and mRNA was isolated using Oligotex (QIAGEN). First-strand cDNA was prepared using random hexamer primers. Control reactions without reverse transcriptase gave no signal. cDNA was used as template in PCR reactions with the following primer sets: p21 5'-CGGTCCCGTGGACAGTGAGC-3' and 5'-AAATCTGT-CAGGCTGCTGCTGCC-3' (amplified fragment size, 371 bp), MDM2 5'-GGAGCGCAAAACGACACTTACA-3' and 5'-CTCGTGTCTGCTGCTGC-TAC-3' (amplified fragment size, 512 bp) and GADD45 5'-AGAGCAGAA-GACCGAAAGGATG-3' and 5'-ACCCGACAGGATGTTGATGTCG-3' (248 bp). Tubulin 5'-GACAGTGTGGCAACCAGATCG-3' and 5'-GTACGGAAGCA-GATGTCGTAG-3' (612 bp) was used as a control.

Northern blot analysis. Cell lines were irradiated with 4 Gy of γ-irradiation and incubated for 4 h before mRNA was prepared. Total RNA was isolated using the TRIZOL (Gibco) method and mRNA was prepared using an Oligotex suspension (Qiagen). Probes for p21, MDM2 and GADD45 were prepared by RT-PCR from mouse RNA and labelled with ³²P-dCTP.

Received 17 February; accepted 21 May 1999.

- Levine, A. J. p53, the cellular gatekeeper for growth and division. *Cell* **88**, 323–331 (1997).
- Prives, C. Signaling to p53: breaking the MDM2–p53 circuit. *Cell* **95**, 5–8 (1998).
- Woo, R. A., McLure, K. G., Lees-Miller, S. P., Rancourt, D. E. & Lee, P. W. DNA-dependent protein kinase acts upstream of p53 in response to DNA damage. *Nature* **394**, 700–704 (1998).
- Smith, G. C. & Jackson, S. P. The DNA-dependent protein kinase. *Genes Dev.* **13**, 916–934 (1999).

- Taccioli, G. E. *et al.* Targeted disruption of the catalytic subunit of the DNA-PK gene in mice confers severe combined immunodeficiency and radiosensitivity. *Immunity* **9**, 355–366 (1998).
- Giaccia, A. J. & Kastan, M. B. The complexity of p53 modulation: emerging patterns from divergent signals. *Genes Dev.* **12**, 2973–2983 (1998).
- Lees-Miller, S. P., Sakaguchi, K., Ullrich, S. J., Appella, E. & Anderson, C. W. Human DNA-activated protein kinase phosphorylates serines 15 and 37 in the amino-terminal transactivation domain of human p53. *Mol. Cell. Biol.* **12**, 5041–5049 (1992).
- Tibbetts, R. S. *et al.* A role for ATR in the DNA damage-induced phosphorylation of p53. *Genes Dev.* **13**, 152–157 (1999).
- Jeggo, P. A. DNA-PK: at the cross-roads of biochemistry and genetics. *Mutat. Res.* **384**, 1–14 (1997).
- Hainaut, P. *et al.* Database of p53 gene somatic mutations in human tumors and cell lines: updated compilation and future prospects. *Nucleic Acids Res.* **25**, 151–157 (1997).
- Rathmell, W. K., Kaufmann, W. K., Hurt, J. C., Byrd, L. L. & Chu, G. DNA-dependent protein kinase is not required for accumulation of p53 or cell cycle arrest after DNA damage. *Cancer Res.* **57**, 68–74 (1997).
- Nussenzweig, A. *et al.* Requirement for Ku80 in growth and immunoglobulin V(D)J recombination. *Nature* **382**, 551–555 (1996).
- Huang, L. C., Clarkin, K. C. & Wahl, G. M. p53-dependent cell cycle arrests are preserved in DNA-activated protein kinase-deficient mouse fibroblasts. *Cancer Res.* **56**, 2940–2944 (1996).
- Gu, Y. *et al.* Growth retardation and leaky SCID phenotype of Ku70-deficient mice. *Immunity* **7**, 653–665 (1997).
- Zhu, C., Bogue, M. A., Lim, D. S., Hasty, P. & Roth, D. B. Ku86-deficient mice exhibit severe combined immunodeficiency and defective processing of V(D)J recombination intermediates. *Cell* **86**, 379–389 (1996).
- Jongmans, W. *et al.* The role of ataxia telangiectasia and the DNA-dependent protein kinase in the p53-mediated cellular response to ionising radiation. *Oncogene* **13**, 1133–1138 (1996).
- Linke, S. P., Clarkin, K. C., Di Leonardo, A., Tsou, A. & Wahl, G. M. A reversible, p53-dependent G0/G1 cell cycle arrest induced by ribonucleotide depletion in the absence of detectable DNA damage. *Genes Dev.* **10**, 934–947 (1996).
- Sakaguchi, K. *et al.* DNA damage activated p53 through a phosphorylation–acetylation cascade. *Genes Dev.* **12**, 2831–2841 (1998).
- McLure, K. G. & Lee, P. W. How p53 binds DNA as a tetramer. *EMBO J.* **17**, 3342–3350 (1998).

Acknowledgements. We thank R. S. Haltiguer and S. Tronick for help with Ab421 coupling. This work was supported by the MRC (A.P., P.A.J.), by a postdoctoral fellowship from the NIH (G.S.J.), by grants from the NIH and from the G. Harold and Leila Y. Mathers Charitable Foundation (G.M.W. and M.B.), in part by the Wellcome Trust, by the Human Frontiers Science Program (E.B.), and in part by the Leukemia Society of America and by a grant from the National Cancer Institute, NIH (G.E.T.). M.H. is an R.M. Phillips senior research fellow; G.E.T. is a scholar of the Leukemia Society of America.

Correspondence and requests for materials should be addressed to G.M.W. (e-mail: wahl@salk.edu). The GenBank accession number for the SCGR11 mutant-p53 sequence is AF151353.

DNA protection by stress-induced biocrystallization

Sharon G. Wolf*, Daphna Frenkiel*, Talmon Arad†, Steven E. Finkel‡, Roberto Kolter‡ & Abraham Minsky*

Departments of * Organic Chemistry and † Structural Biology, The Weizmann Institute of Science, Rehovot 76100, Israel

‡ Department of Microbiology and Molecular Genetics, Harvard Medical School, 200 Longwood Avenue, Boston, Massachusetts 02115, USA

The crystalline state is considered to be incompatible with life. However, in living systems exposed to severe environmental assaults, the sequestration of vital macromolecules in intracellular crystalline assemblies may provide an efficient means for protection. Here we report a generic defence strategy found in *Escherichia coli*, involving co-crystallization of its DNA with the stress-induced protein Dps^{1,2}. We show that when purified Dps and DNA interact, extremely stable crystals form almost instantaneously, within which DNA is sequestered and effectively protected against varied assaults. Crystalline structures with similar lattice spacings are formed in *E. coli* in which Dps is slightly over expressed, as well as in starved wild-type bacteria. Hence, DNA–Dps co-crystallization is proposed to represent a binding mode that provides wide-range protection of DNA by sequestration. The rapid induction and large-scale production of Dps in response to stress, as well as the presence of Dps homologues in many distantly related bacteria, indicate that DNA protection by biocrystallization may be crucial and widespread in prokaryotes.

Under conditions of either oxidative or nutritional stress, *E. coli* produces high levels of the nonspecific DNA-binding protein Dps^{2–4}, which effectively protects DNA against oxidative agents

both *in vitro* and *in vivo*⁴. Dps spontaneously forms spherical dodecamers that are homologous to assemblies formed by ferritins⁵. It has consequently been proposed that the protective activity of Dps stems from its putative ability to sequester iron ions⁵. While effective in attenuating formation of free radicals by the Fenton reaction, a defence strategy based solely on ion sequestration cannot provide the all-encompassing DNA protection that is required—and observed—during prolonged starvation^{1,6}. Indeed, Dps protects DNA against various nucleases² and oxidative agents⁴, thus indicating a general Dps-dependent defence mechanism.

Details of this mechanism were derived from electron microscopy studies of Dps and Dps–DNA complexes (Fig. 1). In the absence of DNA, Dps forms spherical dodecameric particles of ~ 90 Å in diameter², which appear ring-like in projection. Overnight incubation of Dps results in two-dimensional crystals; projection images of these crystals reveal hexagonal packing with a spacing of 78 ± 1 Å, corresponding to a particle diameter of 90 Å (Fig. 1b). An identical value was obtained from the Dps X-ray crystal structure⁵. DNA molecules have a marked effect on Dps organization (Fig. 1c–f). Immediately after the addition of DNA, extensive aggregation of Dps dodecamers occurs that is followed, seconds later, by the

formation of numerous crystals. This rapid crystallization is produced by closed supercoiled plasmids, linear double-stranded DNA and single-stranded RNA molecules, leading in all cases to multi-layered plate-like crystals of similar morphology. No crystals are formed, however, in the presence of negatively charged polypeptides such as polyglutamate.

Although it greatly accelerates crystallization and promotes a three-dimensional morphology, the DNA does not affect the in-plane lattice spacing of the crystals. Irrespective of the DNA structure, this remains at 78 ± 2 Å, similar to that of two-dimensional Dps crystals obtained in the absence of DNA. To assess the effects of the DNA, differential interference contrast (DIC) and fluorescence microscopy studies of DNA–Dps mixtures were carried out. Mixtures of Dps and fluorescently labelled DNA exhibit fluorescent blobs which exclusively coincide with the crystals observed in the DIC mode; crystals obtained with non-labelled DNA are not fluorescent (data not shown). These findings indicate that DNA is incorporated within the crystals and, in conjunction with the electron microscopy results, imply a remarkably efficient nucleic acid–Dps co-crystallization process. The observation that nucleic acids in their various forms do not alter the crystal spacing is

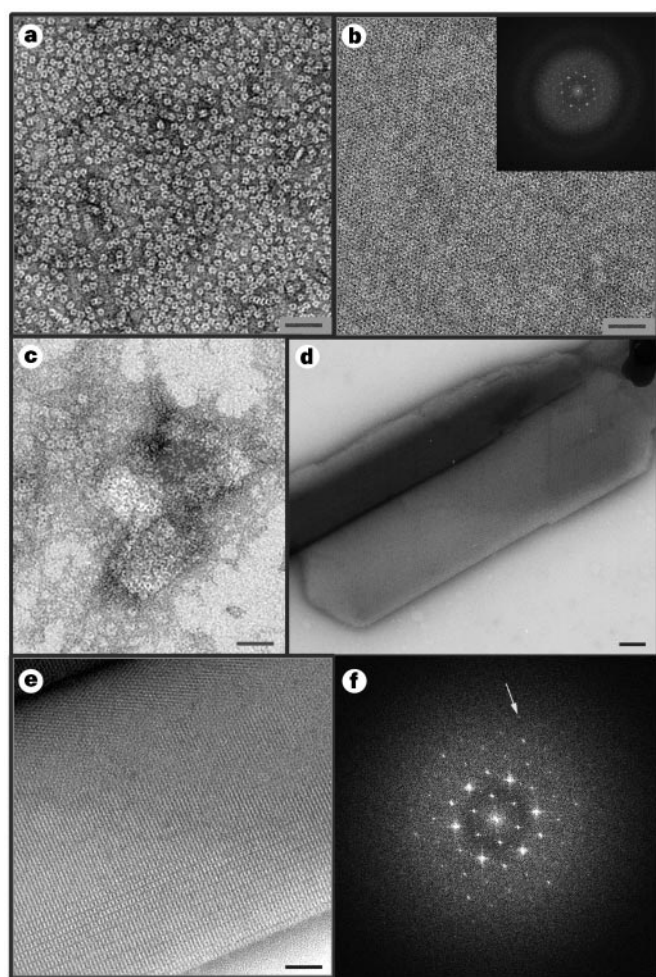


Figure 1 Electron microscopy of Dps and Dps–DNA complexes. **a**, Dps ($70 \mu\text{g ml}^{-1}$). **b**, Dps following overnight incubation. Inset: computed Fourier transform. **c**, **d**, Dps ($50 \mu\text{g ml}^{-1}$) incubated for 2 and 15 s, respectively, with closed circular DNA (Dps/DNA ratio of 1/5, w/w). **e**, Higher magnification of the crystal shown in **d**. **f**, Computed Fourier transform from a Dps–DNA crystal; arrow indicates 18 Å resolution. Scale bars are 40 nm in **a**, **c** and **e**, and 100 nm in **d**.

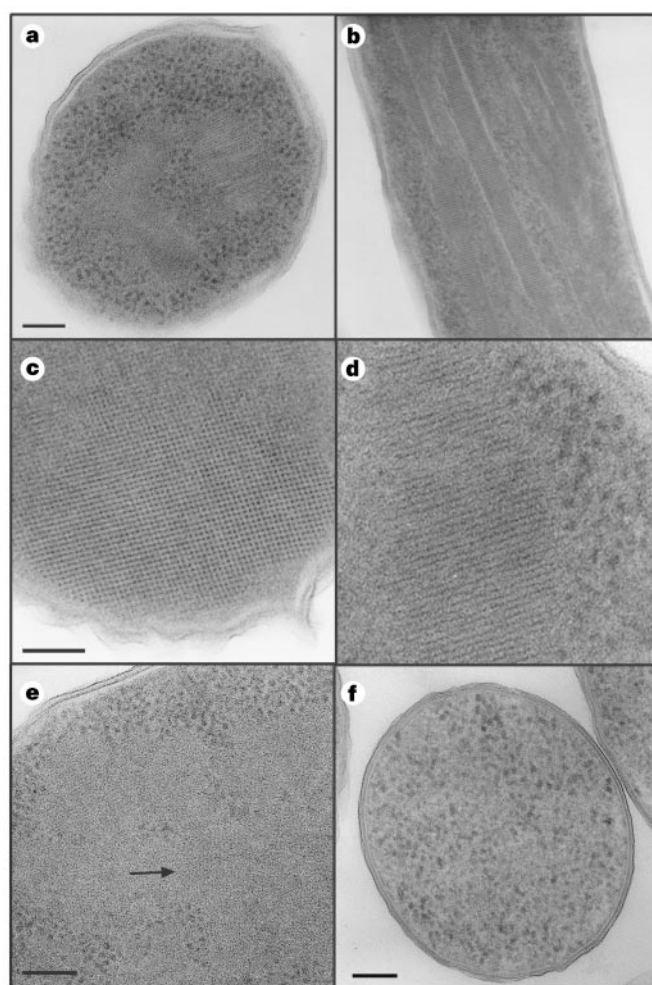


Figure 2 *In situ* DNA–Dps assemblies. **a–d**, Dps-over-producing *E. coli* cells induced at mid-logarithmic phase and then incubated for 48 h. The dark particles that surround the crystals are ribosomes. **e**, Crystalline assemblies in starved wild-type *E. coli*. **f**, Morphology of *dps*[−] bacteria following 48 h of starvation. There are identical crystalline assemblies in bacteria fixed by either cryo or chemical methods. As these fixation modes use different mechanisms, detection of ordered assemblies following both procedures indicates that these structures represent a genuine morphological feature. Scale bars are 100 nm (**a**, **b**, **e**, **f**) and 50 nm (**c**, **d**).

consistent with a model where DNA and Dps form stacked alternating layers, within which DNA is effectively sequestered and protected. This model is reminiscent of DNA-liposome assemblies, where DNA induces a lamellar packing of the lipids and is stacked in between these layers⁷.

Complexes between DNA and nonspecific DNA-binding proteins such as RecA⁸, H-NS⁹, the sporol SASP¹⁰, sperm protamine¹¹ and histone H1 (ref. 12) contain only protein-coated DNA molecules or amorphous aggregates. Thus, the fast DNA-Dps co-crystallization represents an unprecedented binding and protection mode. But is this mode physiologically relevant?

Electron microscopy studies were conducted on *E. coli* that overexpress Dps⁴. Over-production of Dps to levels that are fourfold higher than those accumulated in starved wild-type bacteria results in the appearance of intracellular crystalline structures (Fig. 2a–d). The flat *in vitro* crystals adopt a favoured orientation on the electron microscope grid and hence exhibit only intra-layer spacings. In contrast, information on both intra- and inter-layer spacings is provided by the *in situ* crystalline structures, as they are embedded within randomly sliced sections. Indeed, some of these crystals have a layered morphology, whereas others display square lattices that result from views normal to the layer planes. Although they are smaller and less frequent than the crystals formed following Dps over-production, crystalline assemblies are also detected in starved wild-type *E. coli* (Fig. 2e), thus indicating that ordered assemblies are not merely a consequence of Dps over-expression. The similarity between the *in vitro* and *in vivo* DNA-Dps crystals indicates that the ordered intracellular assemblies are Dps-DNA complexes. Several lines of evidence support this idea. First, *dps* knockout bacteria do not show crystalline assemblies upon starvation (Fig. 2f). Thus, the presence and intracellular amounts of Dps correlate with the appearance, size and frequency of the ordered *in vivo* assemblies. Second, when DNA is isolated from bacteria over-producing Dps, Dps is found to be the predominant protein bound to the DNA. Finally, in actively growing bacteria, chromatin is detected as amorphous spaces¹³. The absence of such regions in bacteria that contain crystalline assemblies implies a co-localization of DNA and the crystals.

Both wild-type and over-producer strains form ordered structures only when starved, when Dps is one of the few proteins still being expressed². This observation points to a causal link between starvation and the crystallization process. Because after prolonged starvation the cellular content of most proteins is reduced¹⁴, any interference that might be imposed by competing DNA-binding proteins upon the ordered DNA-Dps assembly is substantially attenuated.

The stress-related intracellular assembly and the rapid DNA-Dps co-crystallization observed *in vitro* imply a drive for ordered packing that is rarely encountered in living systems or revealed by biopolymers. One intriguing example is the crystalline organization of ferritin in the bacterium *Helicobacter pylori*¹⁵. This organization has been proposed to reduce the risk of iron poisoning by sequestering intracellular iron. The observation that both ferritin and Dps form intracellular crystalline assemblies is provocative in light of their structural homology^{5,16,17} and their abundance in diverse prokaryotes^{17–21}. These findings may indicate an evolutionary pressure to promote biocrystallization processes whose protective activity is exerted through a fast sequestration of damaging agents and their targets. □

Methods

Dps purification. We purified Dps from a Dps over-producer *E. coli* strain (SF101) that harbours plasmids in which the *dps* gene was cloned under the control of the arabinose-inducible *P_{BAD}* promoter⁴. Cells were grown at 37 °C in LB medium to an absorbance at 560 nm of ~1 and induced with 0.3% arabinose. Purification was performed as described², with several modifications. Specifically, we replaced the Sephadex G-100 and Sepharose 6B

columns with a DEAE-cellulose ion-exchange column. The final purification step remained DNA-cellulose affinity chromatography.

DIC and fluorescence microscopy. Dps-DNA complexes (78 µg ml⁻¹ Dps, linearized pBluescript DNA in 10 mM NaCl, 10 mM Tris) were prepared at a Dps/DNA w/w ratio of 1/5. DNA was labelled with YOYO fluorophore before complex formation at a 1/1 molar ratio of dye to base pair (bp). Images in both DIC and fluorescence modes were taken on an Axiovert 35 mm Zeiss microscope with an enhanced Gatan CCD camera.

***In vitro* electron microscopy.** We incubated samples containing Dps and DNA (closed circular supercoiled or linearized pBluescript, 2,958 bp) at concentrations indicated in the caption to Fig. 1 at room temperature in 20 mM NaCl, 10 mM Tris (pH 7.2) on carbon-coated grids. Specimens were stained with 1% uranyl acetate and probed on a Philips CM12 electron microscope with a Tietz CCD, operating at 100 kV.

Thin-section electron microscopy. Wild-type ZK126 *E. coli* cells² were grown in LB medium and starved for 48 h. *E. coli* harbouring the pBAD-*dps* plasmid that carries an arabinose-inducible *dps* gene⁴ were treated with 0.02% arabinose (resulting in Dps accumulation to a level that is 3- to 4-fold higher than that obtained in starved wild-type bacteria). Following induction, we incubated the cells in LB medium at 37 °C for 48 h. Bacteria were fixed by ultra-fast freezing in liquid ethane and subsequently cryosubstituted as described¹³. In control experiments, chemical fixation was performed according to the RK-U procedure²². Samples were embedded in epon; thin sections were stained with 1% uranyl acetate and examined on a Philips CM12 electron microscope.

Received 15 March; accepted 5 May 1999.

- Kolter, R., Siegle, D. A. & Tormo, A. The stationary phase of bacterial life cycle. *Annu. Rev. Microbiol.* **47**, 855–874 (1993).
- Almiron, M., Link, A. J., Furlong, D. & Kolter, R. A novel DNA-binding protein with regulatory and protective roles in starved *E. coli*. *Genes Dev.* **6**, 2646–2654 (1992).
- Altuvia, S., Almiron, M., Huisman, G., Kolter, R. & Storz, G. The *dps* promoter is activated by OxyR during growth and by IHF and sigma S in stationary phase. *Mol. Microbiol.* **13**, 265–272 (1994).
- Martinez, A. & Kolter, R. Protection of DNA during oxidative stress by the nonspecific DNA-binding protein Dps. *J. Bacteriol.* **179**, 5188–5194 (1997).
- Grant, R. A., Filman, D. J., Finkel, S. E., Kolter, R. & Hogle, J. M. The crystal structure of Dps, a ferritin homolog that binds and protects DNA. *Nature Struct. Biol.* **5**, 294–303 (1998).
- Loewen, P. C. & Hengge-Aronis, R. The role of the sigma factor sigma S (KatF) in bacterial global regulation. *Annu. Rev. Microbiol.* **48**, 53–80 (1994).
- Radler, J. O., Koltover, I., Salditt, T. & Safinya, C. R. Structure of DNA-cationic liposome complexes: DNA intercalation in multilamellar membranes in distinct interhelical packing regimes. *Science* **275**, 810–814 (1997).
- Pinsince, J. M. & Griffith, J. D. Early stages in RecA protein-catalyzed pairing. Analysis of coaggregate formation and non-homologous DNA contacts. *J. Mol. Biol.* **228**, 409–420 (1992).
- Tupper, A. E. *et al.* The chromatin-associated protein H-NS alters DNA topology *in vitro*. *EMBO J.* **13**, 258–268 (1994).
- Griffith, J., Makhov, A., Santiago, L. L. & Setlow, P. Electron microscopic studies of the interaction between a *Bacillus subtilis* alpha/beta-type small, acid-soluble spore protein with DNA: protein binding is cooperative, stiffens the DNA, and induces negative supercoiling. *Proc. Natl Acad. Sci. USA* **91**, 8224–8228 (1994).
- Allen, M. J., Bradbury, E. M. & Balhorn, R. AFM analysis of DNA-protamine complexes bound to mica. *Nucleic Acids Res.* **25**, 2221–2226 (1997).
- Bernardin, W. D., Losa, R. & Koller, T. Formation and characterization of soluble complexes of histone H1 with supercoiled DNA. *J. Mol. Biol.* **189**, 503–517 (1986).
- Hobot, J. A. *et al.* Shape and fine structure of nucleoids observed on sections of ultrarapidly frozen and cryosubstituted bacteria. *J. Bacteriol.* **162**, 960–971 (1985).
- Flardh, K., Cohen, P. S. & Kjelleberg, S. Ribosomes exist in large excess over the apparent demand for protein synthesis during carbon starvation in marine *Vibrio* sp. strain CCUG 15956. *J. Bacteriol.* **174**, 6780–6788 (1992).
- Frazier, B. A. *et al.* Paracrystalline inclusions of a novel ferritin containing nonheme iron, produced by the human gastric pathogen *Helicobacter pylori*: evidence for a third class of ferritins. *J. Bacteriol.* **175**, 966–972 (1993).
- Evans, D. J., Evans, D. G., Lampert, H. C. & Nakano, H. Identification of four new prokaryotic bacterioferritins, from *Helicobacter pylori*, *Anabaena variabilis*, *Bacillus subtilis* and *Treponema pallidum*, by analysis of gene sequences. *Gene* **153**, 123–127 (1995).
- Pena, M. M. & Bullerhahn, G. S. The DpsA protein of *Synechococcus* sp. strain PCC7942 is a DNA-binding hemoprotein. Linkage of the Dps and bacterioferritin protein families. *J. Biol. Chem.* **270**, 22478–22482 (1995).
- Pena, M. M., Burkhardt, W. & Bullerhahn, G. S. Purification and characterization of a *Synechococcus* sp. strain PCC 7942 polypeptide structurally similar to the stress-induced Dps/PexB protein of *E. coli*. *Arch. Microbiol.* **163**, 337–344 (1995).
- Chen, L. & Helmman, J. D. *Bacillus subtilis* MrgA is a Dps(PexB) homologue: evidence for metallorepression of an oxidative-stress gene. *Mol. Microbiol.* **18**, 295–300 (1995).
- Bozzi, M. *et al.* A novel non-heme iron-binding ferritin related to the DNA-binding proteins of the Dps family in *Listeria innocua*. *J. Biol. Chem.* **272**, 3259–3265 (1997).
- Wai, S. N., Nakayama, K., Takade, A. & Amako, K. Overproduction of *Campylobacter* ferritin in *E. coli* and induction of paracrystalline inclusion by ferrous compound. *Microbiol. Immunol.* **41**, 461–467 (1997).
- Robinow, C. & Kellenberger, E. The bacterial nucleoid revisited. *Microbiol. Rev.* **58**, 211–232 (1994).

Acknowledgements. This work was supported by a grant from the Consortium of the German Chemical Association. We thank S. Levin-Zidman and E. Shimoni for their help in the electron microscopy studies, and H. Salman for his assistance in the DIC and fluorescence experiments.

Correspondence and requests for materials should be addressed to A.M. (e-mail: cominsky@weizmann.weizmann.ac.il).

Solution structure of the catalytic domain of GCN5 histone acetyltransferase bound to coenzyme A

Yingxi Lin[†], C. Mark Fletcher^{*}, Jianxin Zhou[‡], C. David Allis[‡] & Gerhard Wagner^{*}

^{*} Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, 240 Longwood Avenue, Boston, Massachusetts 02115, USA and MIT/Harvard Center for Magnetic Resonance, Cambridge, Massachusetts 02139, USA

[†] Department of Physics, Harvard University, Cambridge, Massachusetts 02138, USA

[‡] University of Virginia, Department of Biochemistry and Molecular Genetics, 1300 Jefferson Park Avenue, Charlottesville, Virginia 22908, USA

Gene transcription requires the release of inactive DNA from its packaging of histone proteins. Following the discovery of the first transcription-associated histone acetyltransferase, tetrahymena GCN5¹, it was shown that yeast GCN5 is recruited to the promoter and causes hyper-acetylation of histones and transcriptional activation of target genes^{2,3}, establishing a direct connection between histone acetylation and transcriptional activation. Many other important transcription regulators have been found to have histone acetyltransferase activity, including TAF_{II}230/250, p300/CBP and its associated factor PCAF^{4–9}. Here we present the solution structure of the catalytic domain of tGCN5 (residues 47–210) in complex with coenzyme A. The structure contains two domains; the amino-terminal domain is similar to those of other GCN5-related N-acetyltransferases^{10,11} but the carboxy-terminal domain is not. Coenzyme A binds in a deep hydrophobic pocket between the two domains. Chemical shift changes upon titration with histone H3 peptides indicate a binding site at the domain boundary opposite to the coenzyme A site. The structural data indicate a single-step acetyl-transfer reaction mechanism catalysed by a hydrogen bond to the backbone amide group of leucine 126 and the side-chain carboxyl group of a conserved acidic residue.

Proteins of the GCN5 family have four highly conserved sequence motifs (labelled I–IV in Fig. 1a) and a bromodomain that is absent in cytoplasmic histone acetyltransferase (HAT) proteins¹. Deletion mutagenesis has shown that fragments containing motifs II–IV have considerable HAT activity, but all four conserved motifs are required for full activity¹². A related set of sequence motifs has been defined for the GCN5-related N-acetyltransferase (GNAT) superfamily of more than 140 members¹³ (labelled A–D in Fig. 1a). The catalytic domains of the GCN5 family share sequence motifs A, B and D with GNAT superfamily proteins (Fig. 1a). Motif IV, however, is unique to the GCN5 family.

We have determined the solution structure of the tGCN5 catalytic domain (residues 47–210) in complex with coenzyme A (CoA) (Fig. 1). The central seven-stranded β -sheet is composed of two portions, strands S1–S4 and S5–S7, that are internally antiparallel but associate with each other by parallel alignment of their edges. Strand S4 contains a β -bulge in its centre at Ala 124, which divides S4 into two halves, S4' and S4''. A long loop without a regular secondary structure connects strands S6 and S7. This loop exhibits few long-range nuclear Overhauser effects (NOEs) and ¹⁵N-relaxation experiments indicate higher than average mobility (data not shown). This loop contains the end of conserved motif III and most of motif IV (see below).

The structure consists of two domains with the boundary at the end of helix H3 (Fig. 1c). The N-terminal domain forms a contiguous globular fold containing the larger antiparallel β -sheet

(S1–S4) and helices H1 to H3. The smaller C-terminal domain also forms a contiguous globular fold and is connected to the N-terminal domain by the parallel β -sheet between S4' and S5 and side-chain contacts to helix H3.

Coenzyme A is bound in a deep hydrophobic cleft between the two domains. It is enclosed by the end of helix H1, strand S4, the S4–H3 loop and helices H3 and H4. The S4–H3 loop contains the sequence motif Arg/Gln-x-x-Gly-x-Gly/Ala that is conserved in the GNAT family of proteins¹³. The position of CoA in the binding site is defined by 54 NOEs to the protein (Fig. 2) and the conformation of the bound CoA is further defined by 37 intramolecular NOEs.

The interaction of CoA with the protein has three striking features (Fig. 2). First, the pantetheine moiety of CoA resembles a polypeptide chain in which the α -carbons are replaced with ethylene groups. It associates with strand S4'' in a manner reminiscent of an antiparallel β -sheet, with hydrogen bonds to Val 128 HN and Leu 126 CO. In addition, when acetyl-CoA is present, the backbone NH of Leu 126 could hydrogen bond to the oxygen of the acetyl group (this may be important for the catalytic mechanism, as we describe below). Second, the three hydrophobic groups of the pantetheine moiety make contacts with hydrophobic side chains of the protein. The methylene groups at position 2 and 3 interact with Leu 126, Ala 165, Tyr 168 and Phe 169. The methylene groups at positions 6 and 7 interact with Gln 76, Leu 77, Ala 127, Val 128 and Gln 133. The two methyl groups at position 11 and the CH group at position 10 interact with Ala 127, Val 128 and Gln 133. Third, the pyrophosphate group fits into the loop connecting S4 and H3, with hydrogen bonds from the backbone NH groups of three consecutive residues (Val 134, Arg 135 and Gly 136) to three oxygens of the pyrophosphate.

Coenzyme A bends at both the pantetheine moiety and the pyrophosphate group. Compared to the pantetheine group, the adenosine moiety is less well defined (Fig. 1b) and its conformation is determined only by intra-CoA NOEs and a few intermolecular NOEs from Val 134 and Asn 131 (Fig. 2). The base bends back to make hydrophobic contacts with the pantetheine moiety, as shown by NOEs from base proton H8 to positions 10, 11 and 12 of the pantetheine moiety. The 3' phosphate is mostly exposed to the solvent.

To investigate acetyl-CoA binding, we compared the ¹⁵N heteronuclear single quantum coherence (HSQC) spectrum of our existing complex with the spectrum of the complex with acetyl-CoA. There are only a few differences, all involving residues close to the thiol group of the CoA. This indicates that the structures of the two complexes are extremely similar. The β -bulge in strand S4 appears to provide a pocket to accommodate the acetyl group without breaking hydrogen bonds to the adjacent β -strands. Strands S4 and S5 are pulled apart at the bulge, creating an opening of the polypeptide backbone (Fig. 1c). This conformation may provide access for Lys 14 of the histone substrate. The high level of sequence conservation in this region indicates that the β -bulge may be conserved in related proteins (Fig. 1a).

To identify the binding site for the histone substrate, we measured chemical shift changes upon titration of the protein with N-terminal fragments of histone H3 comprising residues 1–21 and 9–19. The results from both peptides are similar. Mapping the chemical shift changes onto the protein structure (Fig. 3a) indicates that the histone substrate binds to a site opposite to the entrance to the CoA-binding pocket, in the cleft between H2 and the long loop connecting S6 and S7. This loop may rearrange upon histone binding to widen the cleft, which would allow Lys 14 of the histone, the target for acetylation, to come closer to the acetyl group on the CoA. The peptide appears to be oriented perpendicular to strands S3 and S4.

Figure 3b shows the electrostatic potential of the binding surface. This reveals a concentration of negative charge provided by six highly conserved acidic residues: Glu 83, Asp 91, Glu 122, Asp 162,

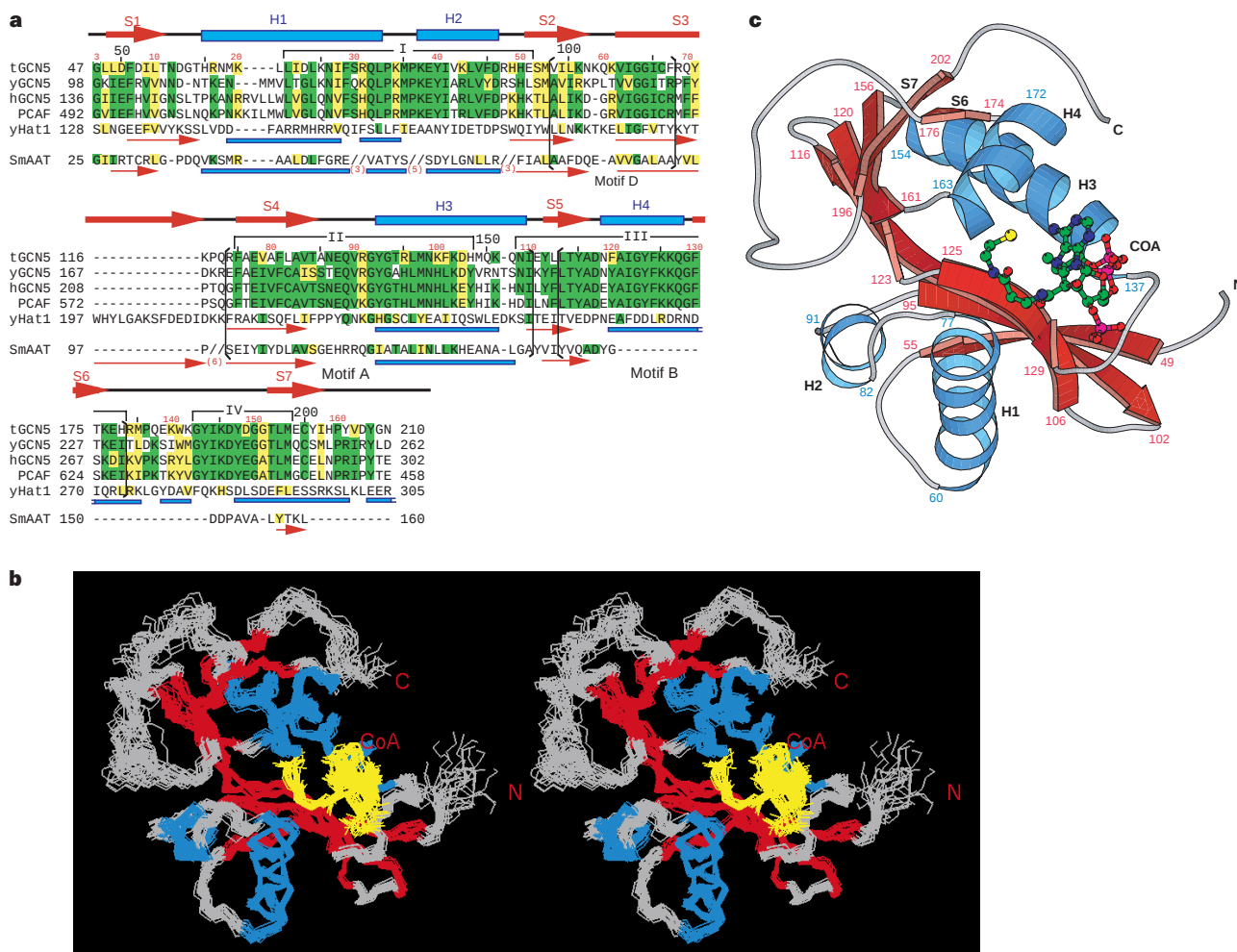


Figure 1 Sequence and structure of the catalytic domain of tGCN5. **a**, Sequence alignment of GCN5 family members with yeast Hat1 protein (yHat1) and *Serratia marcescens* aminoglycoside 3-N-acetyltransferase (SmaAT). Secondary structure elements are shown above the sequences for tGCN5 and below each sequence for yHat1 and SmaAT. For SmaAT, the number of residues shown

in brackets were omitted from the alignment. **b**, Stereo view of 25 calculated structures superimposed on the mean structure. Backbone carbon and nitrogen atoms of the protein and heavy atoms of the CoA are shown. **c**, Ribbon diagram of the structure closest to the mean in the same orientation (produced with Molscript²⁰).

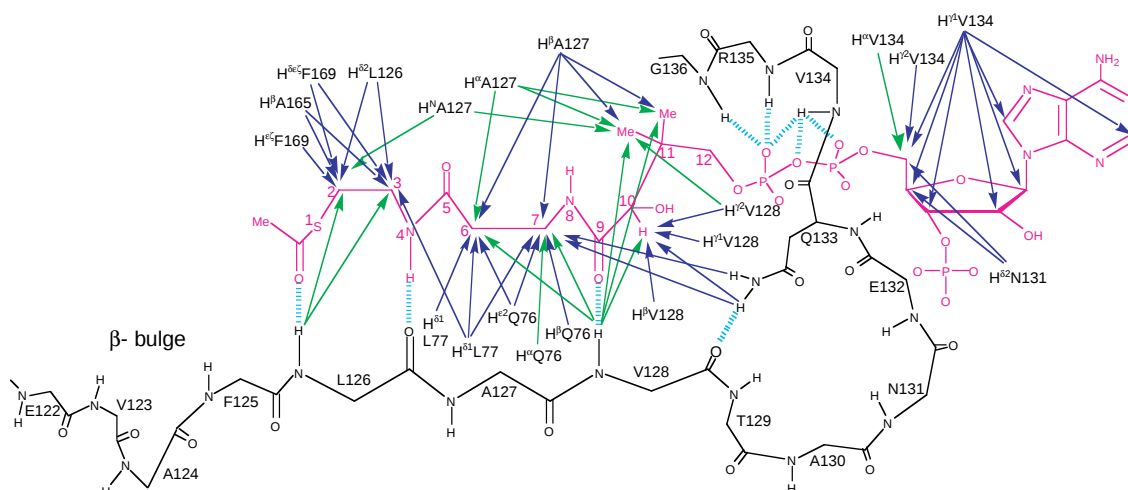


Figure 2 Intermolecular NOEs and hydrogen bonds in the protein-CoA interface. The protein is shown in black and CoA in pink. Blue arrows indicate NOEs between CoA and protein side chains; green arrows indicate NOEs between CoA

and the protein backbone. Note that the acetyl group was not present in the structure we determined and that the hydrogen bond to Leu 126 is hypothetical.

Asp 191 and Asp 193. It seems likely that electrostatic interactions with these residues are important for binding and aligning the lysine-rich histone substrate.

Site-directed mutagenesis of yeast GCN5 has identified many residues that are important for HAT activity^{2,3}. On the basis of the structure and binding site data, we can identify two classes of residues that are sensitive to alanine point mutation. Residues in the first class are most likely to be involved in binding CoA and forming the protein's hydrophobic core (Val 123, Leu 126, Val 128, Gln 133, Arg 135, Gly 136, Tyr 137, Gly 138, Met 142, Tyr 168, Phe 169, Lys 170, Lys 171, Gly 173 and Phe 174). Residues in the second class are more likely to be involved in histone binding and catalysis (His 94, Asp 162, Gly 187, Ile 189, Tyr 192, Asp 193, Thr 196 and Leu 197). Most of the residues in the second group are in the long loop connecting strands S6 and S7, which we have previously described to be part of the histone-binding surface, and are part of conserved motifs III and IV (see above).

Our structural data provide clues about the mechanism of catalysis. The acetyl group could be transferred directly from CoA to Lys 14 of the substrate (single-step mechanism) or indirectly via a covalent intermediate in which the acetyl group is attached to the protein (two-step mechanism). The structural data make the two-step mechanism appear very unlikely, because there are no potential acetyl-group carriers within 10 Å of the CoA's thiol group. Cys 200, which is strictly conserved in GCN5 proteins and is a potential candidate for an acetyl-group carrier, is too far from the CoA (sulphur–sulphur distance of 11.6 ± 0.8 Å) to reach the acetyl group without a large conformational change upon substrate binding, including rearrangement of strand S7.

Therefore, the reaction is likely to occur by a single-step mechanism in which the ϵ -NH₂ group of the substrate lysine makes a nucleophilic attack on the carbonyl carbon of the acetyl group. Catalysis of the single-step mechanism will require polarization of the thioester carbonyl group to enhance the nucleophilic attack and stabilization of the reaction intermediate. The backbone NH group of Leu 126 appears to fulfil these requirements by being positioned to form a hydrogen bond with the carbonyl oxygen of the acetyl group (Fig. 2). Such a hydrogen bond will increase the partial positive charge on the carbonyl carbon, favouring nucleophilic attack by the lysine ϵ -NH₂, and stabilize the negatively charged oxygen of the tetrahedral intermediate.

A second requirement for catalysis of the single-step mechanism is deprotonation of the substrate lysine by a base on the protein. The structure of the tGCN5/CoA complex indicates that the side-chain carboxyl group of Glu 122 or Asp 162 could act as a proton acceptor. Both carboxyl groups are part of the negatively charged patch involved in histone binding (Fig. 3b). The concentration of negative charges in this region may enhance catalysis by favouring neutralization of the catalytic carboxyl group. The carboxyl oxygens of Glu 122 and Asp 162 are, respectively, 15.1 ± 1.8 and 9.7 ± 1.2 Å from the sulphur atom of CoA and would be no more than 12 and 7 Å, respectively, from the acetyl group in the complex with acetyl-CoA. Although these distances seem large, lysine deprotonation does not have to take place in the same site as the acetyl transfer but could occur during substrate binding. In addition, a conformational change upon histone binding could allow either residue to come closer to the acetyl group.

Both Glu 122 and Asp 162 are strictly conserved within the GCN5 family. In addition, mutation to alanine of the residue equivalent to Asp 162 in yeast reduced catalytic activity almost to background level^{2,3}, which, in combination with the structural data, indicates an important role for this residue in catalysis or substrate binding. However, mutation to glutamine of the residue equivalent to Glu 122 in yeast GCN5 causes a drastic decrease in catalytic activity that is much larger at pH 7.5 than at pH 9.75¹⁴, strongly indicating that the role of this residue is deprotonation of the substrate lysine. In addition, a carboxyl group, but not a cysteine or histidine residue,

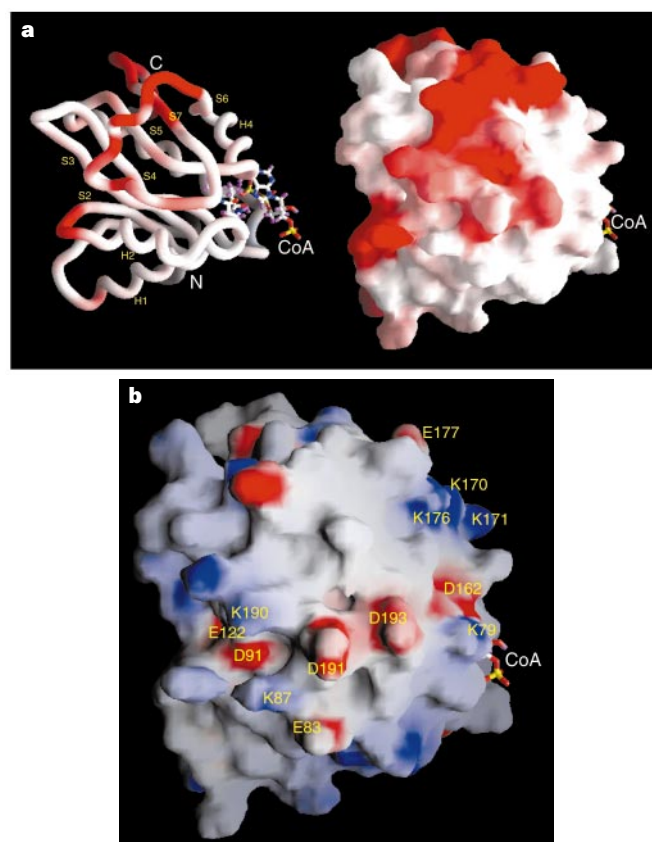


Figure 3 The binding site for histone H3 peptide. **a**, Backbone worm and protein surface showing the histone peptide-binding surfaces in the structure closest to the mean. The colour varies from white to red in proportion to the combined backbone NH and N chemical shift change calculated as $[(\delta_{\text{HN}}^2 + (\delta_{\text{N}}^2/25))/2]^{1/2}$. The molecule has been rotated 20° to the right from the orientation of Fig. 1. **b**, Electrostatic potential of the protein surface in the same orientation. Positive charge is blue and negative is red. Conserved charged residues are labelled. The figure was produced with GRASP²¹.

is required for catalysis by yeast GCN5¹⁴. Thus, the catalytic mechanism we have proposed on the basis of the structural data is strongly supported by mutational and biochemical experiments. Furthermore, a similar mechanism has been proposed for N-myrystoyltransferase¹⁵.

Crystal structures have been reported for two acetyltransferases with limited sequence homology to the GCN5 family and different functions (Fig. 1a). The yeast Hat1 protein (yHat1) acetylates histone H4 (histone H3 is the primary substrate for the GCN5 family) and is believed to be involved in replication-dependent nucleosome assembly rather than transcriptional activation^{11,16}. The *Serratia marcescens* aminoglycoside 3-N-acetyltransferase (SmAAT) acetylates aminoglycoside antibiotics¹⁰. The N-terminal domain of tGCN5 is similar to those of both of these structures, but the C-terminal domain is very different. This is consistent with the role of the C-terminal domain of tGCN5 in substrate binding. In yHat1 the similarity extends as far as the S5 strand and the first part of the H4 helix, producing a similar binding site for the bound acetyl-CoA. There is a backbone hydrogen bond to the carbonyl oxygen of acetyl-CoA in the yHat1 structure, similar to the one we propose for Leu 126 in tGCN5.

In summary, we have presented the three-dimensional structure of tGCN5 with the binding sites for CoA and histone, proposing that charge–charge interactions with conserved acidic residues are important for substrate binding. The structural data are consistent with a single-step acetyl-transfer mechanism and indicate that the backbone NH of Leu 126 and a carboxyl group, probably that of

Glu 122, are involved in catalysis. □

Methods

Sample preparation. The DNA encoding residues 47–210 of tGCN5 was amplified from the cDNA¹ and subcloned into the *NdeI*–*HindIII* site of the pRSET-B vector (Invitrogen). Nine TAA and TAG codons, which encode glutamine in *Tetrahymena*, were mutated to CAA and CAG and verified by DNA sequencing. Phe90 (encoded by TTT) was mistakenly identified as Leu (encoded by CTT) in the sequence published previously¹. The protein was expressed in *Escherichia coli* strain BL21(DE3) in LB or M9 medium with appropriately enriched NH₄Cl and glucose and appropriate D₂O/H₂O ratios. NMR samples contained 0.4–0.5 mM protein in (50 mM NaH₂PO₄/Na₂HPO₄ pH 6.5, 100 mM NaCl, 10 mM DTT, 0.8–1.0 mM unlabelled CoA, 0.1 mM NaN₃, 0.1 mM EDTA).

NMR spectroscopy. NMR spectra were acquired at 27 °C. We obtained protein backbone assignments from triple resonance spectra of 100% ²H-, ¹⁵N- and ¹³C-labelled samples as described¹⁷. H^α protons were assigned from a 3D ¹⁵N TOCSY-HSQC spectrum. Almost complete side-chain resonances were assigned using H(CC)(CO)NH and (H)C(C)(CO)NH spectra of a 65% ²H-, 100% ¹⁵N- and ¹³C-labelled sample and a ¹³C NOESY-HSQC spectrum¹⁹. The CoA protons were assigned from 2D TOCSY and 2D NOESY spectra of a 100% ²H-labelled sample in D₂O. Intramolecular distance restraints were obtained using ¹⁵N NOESY-HSQC, ¹³C NOESY-HSQC and 2D NOESY spectra. Intermolecular restraints were obtained from a 2D ¹⁵N NOESY-HSQC spectrum of a 100% ²H- and ¹⁵N-labelled sample and a 2D ¹³C-edited ¹²C-filtered NOESY spectrum of a 100% ¹⁵N- and ¹³C-labelled sample.

Structure calculation. We derived distance restraints as described¹⁸. There were 2248 intra-protein NOE-derived restraints (1139 intra-residue, 512 sequential, 214 medium range, and 382 long range), 58 intra-protein hydrogen bond restraints for slowly exchanging backbone amide protons in regular secondary structure, 37 intra-CoA restraints and 54 intermolecular restraints. No dihedral angle restraints were used. Structures were calculated using a two-stage protocol¹⁷, adding the CoA in a random conformation and orientation for the second stage. Forty-nine out of 60 starting structures were selected after the first stage and 35 out of 49 structures after the second stage, on the basis of small total X-PLOR energies. An additional 10 structures were discarded because the CoA adenine group had a conformation that was not consistent with the NOE data. The final 25 structures have no NOE violations greater than 0.4 Å, and 93.7% of the non-glycine and non-proline residues lie in the most favourable and additionally allowed regions of the Ramachandran plot. The mean r.m.s. deviation from the mean structure is 0.54 ± 0.102 Å for the backbone heavy atoms of residues 49–176 and 196–203, 1.10 ± 0.108 Å for all heavy atoms of the same residues, and 0.59 ± 0.12 Å for all heavy atoms of CoA. The r.m.s. deviations from ideal geometry are 0.0030 Å for bonds, 0.48° for bond angles and 0.34° for improper angles. Figures were produced with InsightII program (Molecular Simulations Inc., San Diego) (Fig. 1b), Molscript²⁰ (Fig. 1c), and GRASP²¹ (Fig. 3).

Histone peptide titrations. We performed histone peptide titrations with 0.25-mM samples of ¹⁵N-labelled protein, recording ¹⁵N HSQC spectra before and after addition of the unlabelled histone H3 peptides (1)-ARTKQTARK-STGGKAPRKQLQ-(21) and (9)-KSTGGKAPRKQ-(19) at concentrations of up to 1.3 and 1.0 mM, respectively.

Received 16 February; accepted 12 May 1999.

1. Brownell, J. E. *et al.* Tetrahymena histone acetyltransferase A: a homolog to yeast Gcn5p linking histone acetylation to gene activation. *Cell* **84**, 843–851 (1996).
2. Kuo, M. H., Zhou, J., Jambeck, P., Churchill, M. E. A. & Allis, C. D. Histone acetyltransferase activity of yeast Gcn5p is required for the activation of target genes *in vivo*. *Genes Dev.* **12**, 627–639 (1998).
3. Wang, L. *et al.* Histone acetyltransferase activity is conserved between yeast and human GCN5 and is required for complementation of growth and transcriptional activation. *Mol. Cell. Biol.* **17**, 519–527 (1997).
4. Bannister, A. J. & Kouzarides, T. The CBP co-activator is a histone acetyltransferase. *Nature* **384**, 641–643 (1996).
5. Grant, P. A. *et al.* A subset of TAF(II)s are integral components of the SAGA complex required for nucleosome acetylation and transcriptional stimulation. *Cell* **94**, 45–53 (1998).
6. Mizzen, C. A. *et al.* The TAF(II)250 subunit of TFIID has histone acetyltransferase activity. *Cell* **87**, 1261–1270 (1996).
7. Ogryzko, V. V., Schiltz, R. L., Russanova, V., Howard, B. H. & Nakatani, Y. The transcriptional coactivators p300 and CBP are histone acetyltransferases. *Cell* **87**, 953–959 (1996).
8. Ogryzko, V. V. *et al.* Histone-like TAFs within the PCAF histone acetylase complex. *Cell* **94**, 35–44 (1998).
9. Yang, X. J., Ogryzko, V. V., Nishikawa, J., Howard, B. H. & Nakatani, Y. A p300/CBP-associated factor that competes with the adenoviral oncoprotein E1A. *Nature* **382**, 319–324 (1996).

10. Wolf, E. *et al.* Crystal structure of a GCN5-related N-acetyltransferase: *Serratia marcescens* aminoglycoside 3-N-acetyltransferase. *Cell* **94**, 439–449 (1998).
11. Dutnall, R. N., Tafrov, S. T., Sternglanz, R. & Ramakrishnan, V. Structure of the histone acetyltransferase Hat1: a paradigm for the GCN5-related N-acetyltransferase superfamily. *Cell* **94**, 427–438 (1998).
12. Candau, R., Zhou, J. X., Allis, C. D. & Berger, S. L. Histone acetyltransferase activity and interaction with ADA2 are critical for GCN5 function *in vivo*. *EMBO J.* **16**, 555–565 (1997).
13. Neuwald, A. F. & Landsman, D. GCN5-related histone N-acetyltransferases belong to a diverse superfamily that includes the yeast SPT10 protein. *Trends Biochem. Sci.* **22**, 154–155 (1997).
14. Tanner, K. G. *et al.* Catalytic mechanism and function of invariant glutamic acid-173 from the histone acetyltransferase GCN5 transcriptional activator. *J. Biol. Chem.* **274**, 18157–18160 (1999).
15. Bhatnagar, R. S. *et al.* Structure of N-myristoyltransferase with bound myristoyl-CoA and peptide substrate analogs. *Nature Struct. Biol.* **5**, 1091–1097 (1998).
16. Parthun, M. R., Widom, J. & Gottschling, D. E. The major cytoplasmic histone acetyltransferase in yeast: links to chromatin replication and histone metabolism. *Cell* **87**, 85–94 (1996).
17. Matsuo, H. *et al.* Structure of translation factor eIF4E bound to m⁷GDP and interaction with 4E-binding protein. *Nature Struct. Biol.* **4**, 717–724 (1997).
18. Fletcher, C. M., Pestova, T. V., Hellen, C. U. T. & Wagner, G. Structure and interactions of the translation initiation factor eIF1. *EMBO J.* **18**, 2631–2637 (1999).
19. Lin, Y. & Wagner, G. Efficient side chain and backbone assignments in large proteins. *J. Am. Chem. Soc.* (submitted).
20. Kraulis, P. J. MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallogr.* **24**, 946–950 (1991).
21. Nicholls, A., Sharp, K. A. & Honig, B. Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. *Proteins* **11**, 281–296 (1991).

Acknowledgements. We thank R. Marmorstein for initiating this work and providing the histone peptides; G. Heffron for help with NMR experiments; V. Ramakrishnan for providing the coordinates of yHat1 before their release by the PDB; and J. Denu for providing a copy of his paper before publication. This project was supported by grants from NSF, NIH, the Harvard Center for Structural Biology and the Giovanni Armenise-Harvard Foundation for Advanced Scientific Research.

Correspondence and requests for materials should be addressed to G.W. (e-mail: wagner@wagner.med.harvard.edu). Atomic coordinates have been deposited in the Protein Data Bank, accession code 5GCN.

erratum

The CED-4-homologous protein FLASH is involved in Fas-mediated activation of caspase-8 during apoptosis

Yuzuru Imai, Takaharu Kimura, Akira Murakami, Nobuyuki Yajima, Kazuhiro Sakamaki & Shin Yonehara

Nature **398**, 777–785 (1999)

In this Article, the GenBank accession number of the nucleotide sequence of mouse FLASH was deleted. The accession number is AF132726. □

correction

Structural basis for initiation of transcription from an RNA polymerase–promoter complex

Graham M. T. Cheetham, David Jeruzalmi & Thomas A. Steitz

Nature **399**, 80–83 (1999)

The accession number of the coordinates deposited in the Protein Data Bank was wrongly given as 1CIZ: the correct accession number is 1CEZ. □

Going down the tubes

Infrared analysers are flavour of the month in the food and pharmaceuticals industries, and a recent patent seeks to replace the familiar test tube. All this and alcohol too.

Test tube Mark 2

From Norwind-Cortez

A new slant on the test tube

The familiar test tube will be old hat, if this newly patented variant finds favour in the laboratory. Prototypes (pictured right) have been produced by the biotech company Norwind-Cortez, and the company is negotiating licensing agreements on the tube. One of its claimed advantages is the fact that it can be used in either the prone or vertical position, so is not dependent on a vacant slot in a test-tube rack. In the prone position the tube's contents can be heated or cooled more quickly than in the traditional tube. In cell culture, the prone tube offers more contact with the air for a given volume, and better dispersion of the cell culture. Contact number for the Michigan-based company is (734) 528 1099.

Reader Service No. 100



Prone to standing up

are savouring fruit juices, still drinks and milk.

Reader Service No. 101

Pipette Tracker

From Alpha Laboratories www.alphalabs.co.uk

A free download to track pipettes

A free copy of Alpha Pipette Tracker can now be downloaded from the Internet. Pipette Tracker is a software tool which monitors, controls, records and reports the performance of liquid measuring devices and balances. The system is suitable for pipettes, burettes, bottle-top dispensers, volumetric glassware and robotic liquid-handling systems.

Reader Service No. 102

AlcoTrace

From Trace

www.trace-ag.de

If you've got the bottle, this will measure the alcohol

AlcoTrace, made by TRACE Biotech AG, is designed to determine alcohol content without the need for sample preparation, dilution or degassing. Simply open the bottle, push the button, and wait just a short time to get the result. AlcoTrace does this by combining distillate separation and enzymatic biosensor technologies. There are three measurement ranges to cater for all palates: 0–1 vol.% for non-alcoholic soft drinks and non-alcoholic beer, 0–7 vol.% for beer, and 0–18 vol.% for wines.

Reader Service No. 103

Soil analyser

From ADC BioScientific Ltd.

The answer lies in the soil

ADC has introduced a new fully integrated system for the analysis of soil respiration, in multiple samples, for agrochemical registration and ecological research. Investigations on microbial biochemical activity are achieved by measuring the evolution of carbon dioxide from a soil sample. From this data, rates of respiration and amounts of



Down-to-earth breath test

microbial biomass present can be determined. Incorporating a high-performance carbon dioxide infrared gas analyser the ADC Soil Respiration Measurement System combines accurate and reliable gas analysis with versatile experimental programming. Up to 24 soil samples may be sequentially measured by the system that features a full-colour graphic display, flow control and internal and remote data-storage facilities. Optional sensors for determining soil moisture and temperature can be connected directly to the system.

Reader Service No. 104

Brewer's Assistant

From Perkin Elmer

www.perkin-elmer.com

Near-infrared analysers are getting cleverer

In the manufacturer's words, the Brewer's Assistant "turns brewery production floors into satellite laboratories". It makes use of the latest refinements in near-infrared analysis. So the Brewer's Assistant contains all that is required to measure alcohol by volume, original gravity, real and apparent extract, and calories. The device can be satisfied with just 5 ml of filtered beer, and can analyse up to 50 beers per hour.

Reader Service No. 105

LabSpec Pro

From Analytical Spectral Devices, Inc.

www.asdi.com

A spectrophotometer that travels

ASD's LabSpec Pro is a portable near-infrared analyser for the chemical, pharmaceutical, petrochemical and food industries. It weighs only 15 pounds and can be fitted with a range of probes and sampling accessories for the fibre-optic input to measure liquids, powders and slurries. Its wavelength range from 350 nm to 2,500 nm can penetrate plastic or glass, so samples can be tested in their original containers. LabSpec Pro measures 15 × 4.4 × 12.5 inches and oper-

Electronic organs

From NST

www.nordicsensor.com

Second generation 'electronic nose'

The NST 3320 is the second generation 'electronic nose' from Nordic Sensor Technologies. Its all-in-one design incorporates a robotic needle and an auto-sampler. The latest version of the nose offers increased sensitivity and accuracy for detecting odours in samples with very low volatility. This is achieved by using a heated sample turntable to speed up the release of odour from the test material. Applications include classification of emissions, detection of odourless emissions and analysis of toxic, allergenic or carcinogenic emission. NST say the next step for their sensor technology is the Liquid Analyser, or 'electronic tongue'. Prototypes



Do you have a nose for a bargain?

ates on AC/DC or a NiMH battery. An optional carrying case incorporates space for a laptop computer and accessories.
Reader Service No. 106

Nepheloskan Ascent

From LabSystems www.labsystems.fi
Rapid particle measurement

LabSystems claim that the Nepheloskan Ascent is the first commercially available nephelometer for measuring particles in microplate format. The nephelometer offers greater sensitivity than photometers, and no interference from a coloured measurement medium. Depending on the measurement type, a 96-well plate can be read in just 25 seconds. Such sensitivity allows the testing of particles for applications such as drug solubility tests, quantitation of specific proteins and antibiotic sensitivity studies.
Reader Service No. 107

Micropositioner

Charles Supper Company
www.charles-supper.com

A fully adjustable micropositioner

This compact micropositioner is fully adjustable, with a lockable x-y base and concentrically stacked arcs to permit fine-tune positioning. The Supper ultraPositioner is



Get in position

machined from aluminium and black-anodized with a non-reflective finish. Total x-y travel is 0.25 inches, across a 30° cross-axis arc.
Reader Service No. 108

SPE wall chart

From Varian www.varian.com
The writing is on the wall

Varian's Sample Preparation Products wall chart features SPE method development tools and tips. The 3-foot × 5-foot full-colour poster illustrates many SPE terms normally listed in reference books and charts such as method troubleshooting and selecting the best sorbents and solvents. Sorbent, matrix, and analyte properties are also highlighted. Some 54 acid/base properties (pKats) of various functional groups and common SPE terms are listed.
Reader Service No. 109

pH sensor

From Select Systems www.electrode.co.uk
A miniature stainless steel pH sensor for microplates

The probe of this pH sensor stores dry and requires no maintenance, unusual for a pH electrode designed for the measurement of small volumes. This makes it well suited for performing measurements in micro plates, small test tubes and vials. Temperature compensation is automatic or manual. Up to three buffers can be recognized automatically and the operator can choose between one or two point calibration. Depending on whether speed or accuracy is most important, the pH resolution can be adjusted from 0.01 to 0.1 pH.

Reader Service No. 110

Isolute-96

From International Sorbent Technology
www.ist-spe.com

A 96-well format solid-phase extraction plate

Isolute-96 was developed to speed up solid-

phase extraction (SPE) method development. Application development plates are available for basic drugs (containing six cation exchange and mixed-mode sorbents), acidic drugs (containing anion exchange and mixed-mode sorbents) and nonpolar/basic drugs (containing six non-polar non-endcapped sorbents, endcapped C18 and ENV+). Isolute-96 plates are supplied with 50 mg of sorbent per well. Using the Isolute-96 application development plates enables a range of suitable sorbents to be screened simultaneously to optimize recovery and extract purity. Technical notes for use with the Isolute sorbents are also available, describing general approaches to the use of mixed-mode and non-polar, non-endcapped sorbents for the extraction of basic drugs.

Reader Service No. 111

Top Count

From Packard Instrument Co.
www.packardinst.com

You can count on this

Top Count NXT microplate scintillation and luminescence counters can automatically calculate and report per cent lysis, for cytotoxicity assays, or per cent inhibition for binding or enzyme assays. The multi-group sample mapping feature allows totals, blanks and their corresponding unknown samples to be mapped in virtually any configuration desired. After counting, the totals, blanks and their corresponding unknowns are automatically linked to calculate and report results. The advantage of this feature is that it simplifies the preparation of samples, and it overcomes the need for exporting the raw data to perform the calculations.

Reader Service No. 112

These notes are compiled in the Nature office from information provided by the manufacturers. For more details, fill in the reader service card bound inside the journal.

ADVERTISEMENTS

luminometer.committed™

luminometer.com

A website of luminometers and their application methods.

TURNER DESIGNS
845 W. Maude Avenue • Sunnyvale, CA 94086
(408) 749-0994 • FAX (408) 749-0998

READER ENQUIRY NO. 10

Trichostatin A

CC1=CC=C(C=C1)C(=O)C(C)C(=C)C(=O)N

from *Streptomyces*
Catalog No. 204-11991

This histone deacetylase inhibitor is highly effective at low concentrations, with high specificity. For more information, please call
Wako BioProducts: (800) 992-9256; (804) 271-7677.

READER ENQUIRY NO. 98

Homogenizers?

- Tissue-Tearor™
- BeadBeater™
- Mini-BeadBeater™
- BioHomogenizer™

Check Us Out

www.biospec.com

BioSpec Products, Inc.
Ph/Fax 918-336-3363
e-mail info@biospec.com

New Probe Sizes
0.45cm!

READER ENQUIRY NO. 54